

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 16:39:17 ; Search time 27.0743 Seconds
(without alignments)
1559.306 Million cell updates/sec

Title: US-09-902-563-2
Perfect score: 2378
Sequence: 1 MKLANWYWLSSAVLATYGF.....GYKSSFKEAKYMIRPKHKFP 439

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2378	100.0	439	137391	fibrinogen-like pr
2	1853.5	77.9	432	A27447	cytotoxic T-lympho
3	1849.5	77.8	432	I56934	fibrinogen-like pr
4	511.5	21.5	463	A38463	fibrinogen beta ch
5	506.5	21.3	312	JN0596	fibrinogen-related
6	484.5	20.4	453	FGHUGB	fibrinogen gamma-B
7	480.5	20.2	437	FGHUG	fibrinogen gamma-A
8	479.5	20.2	479	A25052	fibrinogen beta ch
9	470	19.8	468	FGBOB	fibrinogen beta ch
10	462	19.4	444	S05313	fibrinogen gamma-B
11	461	19.4	491	PGHJB	fibrinogen beta ch
12	456	19.2	438	A32670	fibrinogen gamma c
13	452	19.0	334	JC5980	ficolin-A precursor
14	451	19.0	866	D44234	fibrinogen alpha c
15	446	18.8	282	A35084	fibrinogen-related
16	434.5	18.3	326	B47172	ficolin-beta - pig
17	431	18.1	328	A05299	fibrinogen beta ch
18	428.5	18.0	323	A47172	transforming growt
19	428.5	18.0	432	FGLMGS	fibrinogen gamma c
20	423	17.8	326	S61517	ficolin-1 precursor
21	423	17.8	1353	JH0675	restrictin precurs
22	417.5	17.6	1356	A45445	janusin precursor,
23	404	17.0	437	FGRTGA	fibrinogen gamma-A
24	404	17.0	445	FGRTGB	fibrinogen gamma-B
25	403.5	17.0	220	S28170	tenascin homolog -
26	403.5	17.0	417	S65944	tenascin-X - pig (
27	401.5	16.9	4135	T43629	tenascin-X - bovin
28	397.5	16.7	3566	A40701	tenascin-X precurs
29	396	16.7	2019	JQ1322	tenascin precursor

30	393.5	16.5	860	2	I48839	tenascin-X - mouse
31	393.5	16.5	4006	2	T09070	probable tenascin
32	386	16.3	2201	2	A32160	tenascin-C - human
33	386.5	16.3	1746	1	S19694	tenascin precursor
34	386	16.2	1810	1	A32230	tenascin precursor
35	382.5	16.1	1914	2	T42635	tenascin Y precurs
36	356.5	15.0	641	1	A41932	fibrinogen alpha-1
37	346.5	14.6	774	2	A39832	scabrous locus (sc
38	215.5	9.1	463	2	T15876	hypothetical prote
39	198	8.3	915	2	T21773	hypothetical prote
40	198	8.3	927	2	T21772	hypothetical prote
41	184.5	7.8	127	2	PC2036	microfibril-associ
42	160	6.7	146	2	T32255	hypothetical prote
43	160	6.7	431	2	T29850	hypothetical prote
44	160	6.7	452	2	T26827	hypothetical prote
45	140	5.9	933	2	A31930	cytotactin - chick

ALIGNMENTS

RESULT 1

I37391
fibrinogen-like protein expressed in T lymphocytes (pT43) - human
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000
C;Accession: I37391; S47273
R;Ruegg, C.; Pytela, R.
Gene 160, 257-262, 1995
A;Title: Sequence of a human transcript expressed in T-lymphocytes and encoding a fibrinogen gamma chain; fibrinogen beta/gamma homology
A;Reference number: I37391; MJID:95369700; PMID:7642106
A;Accession: I37391
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-439 <RES>
A;Cross-references: EMBL:Z36531; NID:G535184; PIDN:CAA85298.1; PID:G535185
A;Note: submitted to the EMBL Data Library, August 1994
C;Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
F;210-435/Domain: fibrinogen beta/gamma homology <FBG>

Query Match: 100.0%; Score 2378; DB 2; Length 439;
Best local Similarity 100.0%; Pred. No. 4e-157;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY	1	MKLANWYWLSSAVLATYGFVLVANNETEEIKDERAKDVCPVRLESRGKCEAGECPYQVS	60
EB	1	MKLANWYWLSSAVLATYGFVLVANNETEEIKDERAKDVCPVRLESRGKCEAGECPYQVS	60
QY	61	LPPLTIQLPKQFSRIEEVFKEVONLKEIVNSLKKSCQCKLQADDNGCPGRNGLLPSTG	120
DB	61	LPPLTIQLPKQFSRIEEVFKEVONLKEIVNSLKKSCQCKLQADDNGCPGRNGLLPSTG	120
QY	121	APGEVGDNRVRELESEVNKLSELKNAKEEINVHGRLEKLNVMNNIENYVDSKVANL	180
DB	121	APGEVGDNRVRELESEVNKLSELKNAKEEINVHGRLEKLNVMNNIENYVDSKVANL	180
QY	181	TFVNSLDGKCKSPSQEQIQSRPVQHLIYKDCSDYYAIGKRSSSETYRVTDPKNSSEFV	240
DB	181	TFVNSLDGKCKSPSQEQIQSRPVQHLIYKDCSDYYAIGKRSSSETYRVTDPKNSSEFV	240
QY	241	YCDMETMGGWTVLQARLDGSTNFTRTWDYKAGFGNLRREFWLGNDKIHLTKSKEMIL	300
DB	241	YCDMETMGGWTVLQARLDGSTNFTRTWDYKAGFGNLRREFWLGNDKIHLTKSKEMIL	300
QY	301	RIDLEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTPD	360
DB	301	RIDLEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTPD	360
QY	361	KNDRYPSGNCGLYSSGWFWFDCLSANLNGKYYHQYKGVGRNGIFWGTWPGVSEAHPPG	420
DB	361	KNDRYPSGNCGLYSSGWFWFDCLSANLNGKYYHQYKGVGRNGIFWGTWPGVSEAHPPG	420
QY	421	YKSSFKEAKMIRPKHKP	439

Db 421 YKSSFKEAKMMIRPKHKFP 439

RESULT 2

A27447

Cytotoxic T-lymphocyte-specific protein precursor (clone pT49) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 03-Dec-1999

C;Accession: A27447

R;Koyama, T.; Hall, L.R.; Haser, W.G.; Tonegawa, S.; Saito, H. Proc. Natl. Acad. Sci. U.S.A. 84, 1609-1613, 1987

A;Title: Structure of a cytotoxic T-lymphocyte-specific gene shows a strong homology to

A;Reference number: A27447; MUID:87175527; PMID:3550794

A;Accession: A27447

A;Molecule type: mRNA

A;Residues: 1-432 <KC>

A;Cross-references: GB:M16238; NID:G193304; PIDN:AAA37624.1; PID:G387156

C;Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology

F;203-428/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 77.9%; Score 1853.5; DB 2; Length 432;

Best Local Similarity 77.7%; Pred. No. 8.1e-121;

Matches 341; Conservative 42; Mismatches 49; Indels 7; Gaps 4;

QY 1 MKLANWYWLSSAVLATYGFVAVANNETEIEIKDERAKDVCPVRLESRGKCEBAGECPYQVS 60

Db 1 MRLPGWLWSSAVLAACR-AVEEHNLTGLEDASAAQACPARLEGSGRC-EGSQCPQLT 59

QY 61 LPPLTIQLPKQPSRIEEVFKEVQNLKEIVNSLKKSCQDCKLOADDNGDPGRNGLLPSTG 120

Db 59 LPTLTIQ:PROLGSMEEVLTKEAVDSLKKSCQDCKLOADDHROPGGNG---GNG 114

QY 121 APGEVGNRVRELESEVNVKLSSELKNKAKEEINVLHGRLKLNLMNMNIENYVDSKYANL 180

Db 115 AE-TAEDSRVQCELESQVNVKLSSELKNKAKDQIQGLQGRLETLHLVNMNIENYVDNKVANL 173

QY 181 TFVNSLDGKCKSKCPSQEQIQSRPVQHLIYKCCSDYVAIGKRSSETYRVTPDPKNSFEV 240

Db 174 TVVNSLDGKCKSKCPSQEHMQSQPVQHLIYKCCSDHYVLGRSSGAYRVTPDHRNSSFEV 233

QY 241 YCDMETMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLTKSKEMIL 300

Db 234 YCDMETMGGGWTVLQARLDGSTNFTREWQDYKAGFNLRRREFWLGNDKIHLTKSKEMIL 293

QY 301 RIDLEDENGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTPD 360

Db 294 RIDLEDENGTLTYALYDQFYVANEFLKYRLHIGNYNGTAGDALRFNKHYNHDLRFFTPD 353

QY 361 KDNDRYPSGNCGLYYSSGWWFDCLSANLNGKYYHQYRGVRNGIFWGTWPGVSEAHPGG 420

Db 354 RDNDRYPSGNCGLYYSSGWWFDCLSANLNGKYYHQYRGVRNGIFWGTWPGINQAQPGG 413

QY 421 YKSSFKEAKMMIRPKHKFP 439

Db 414 YKSSFQAKMMIRPKHKFP 432

RESULT 3

I56934

fibrinogen-like protein - mouse

C;Species: Mus sp. (mouse)

C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 03-Dec-1999

C;Accession: I56934

R;Parr, R.L.; Fung, L.; Reneker, J.; Myers-Mason, N.; Leibowitz, J.L.; Levy, G. J. Virol. 69, 5033-5038, 1995

A;Title: Association of mouse fibrinogen-like protein with murine hepatitis virus-induced

A;Reference number: I56934; MUID:95333285; PMID:7609373

A;Accession: I56934

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-432 <RES>

A;Cross-references: GB:S78773; NID:G1042169; PIDN:AAE34823.1; PID:G1042170

C;Genetics:

A;Gene: mustiblp

C;Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology

F;203-428/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 77.8%; Score 1849.5; DB 2; Length 432;

Best Local Similarity 77.4%; Pred. No. 1.5e-120;

Matches 340; Conservative 42; Mismatches 50; Indels 7; Gaps 4;

QY 1 MKLANWYWLSSAVLATYGFVAVANNETEIEIKDERAKDVCPVRLESRGKCEBAGECPYQVS 60

Db 1 MRLPGWLWSSAVLAACR-AVEEHNLTGLEDASAAQACPARLEGSGRC-EGSQCPQLT 59

QY 61 LPPLTIQLPKQPSRIEEVFKEVQNLKEIVNSLKKSCQDCKLOADDNGDPGRNGLLPSTG 120

Db 59 LPTLTIQ:PROLGSMEEVLTKEAVDSLKKSCQDCKLOADDHROPGGNG---GNG 114

QY 121 APGEVGNRVRELESEVNVKLSSELKNKAKEEINVLHGRLKLNLMNMNIENYVDSKYANL 180

Db 115 AE-TAEDSRVQCELESQVNVKLSSELKNKAKDQIQGLQGRLETLHLVNMNIENYVDNKVANL 173

QY 181 TFVNSLDGKCKSKCPSQEQIQSRPVQHLIYKCCSDYVAIGKRSSETYRVTPDPKNSFEV 240

Db 174 TVVNSLDGKCKSKCPSQEHMQSQPVQHLIYKCCSDHYVLGRSSGAYRVTPDHRNSSFEV 233

QY 241 YCDMETMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLTKSKEMIL 300

Db 234 YCDMETMGGGWTVLQARLDGSTNFTREWQDYKAGFNLRRREFWLGNDKIHLTKSKEMIL 293

QY 301 RIDLEDENGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTPD 360

Db 294 RIDLEDENGTLTYALYDQFYVANEFLKYRLHIGNYNGTAGDALRFNKHYNHDLRFFTPD 353

QY 361 KDNDRYPSGNCGLYYSSGWWFDCLSANLNGKYYHQYRGVRNGIFWGTWPGVSEAHPGG 420

Db 354 RDNDRYPSGNCGLYYSSGWWFDCLSANLNGKYYHQYRGVRNGIFWGTWPGINQAQPGG 413

QY 421 YKSSFKEAKMMIRPKHKFP 439

Db 414 YKSSFQAKMMIRPKHKFP 432

RESULT 4

A38463

fibrinogen beta chain - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 13-Aug-1999

C;Accession: A38463

R;Weissbach, J.; Oddoux, C.; Procyk, R.; Grieninger, G. Biochemistry 30, 3290-3294, 1991

A;Title: The beta chain of chicken fibrinogen contains an atypical thrombin cleavage s

A;Reference number: A38463; MUID:91182745; PMID:2009266

A;Accession: A38463

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-463 <WEI>

A;Cross-references: GB:M58514; NID:G211779; PIDN:AAA48770.1; PID:G211780

C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disul

F;773-202/Domain: fibrinogen disulfide ring homology <FDR>

F;212-460/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 21.5%; Score 511.5; DB 2; Length 463;

Best Local Similarity 30.4%; Pred. No. 7.7e-28;

Matches 137; Conservative 71; Mismatches 154; Indels 89; Gaps 16;

QY 31 KDERAKDVCPVRLESRGK----EEAGE-CPYQVSLPPLTIQLPKQPSRIEEVFKEVQN- 84

Db 51 QDKQAMKKGPIIYDPDAGGCKHPLDELGVLCPTGCE---LQTTLLKQEKTVKPVLRDLKDR 107

QY 85 -----LKEIVNSLKKSCQDCKLOADDNGDPGRNGLLPSTGAPGEVGNRVRELE 134

Db 108 VAKFSDTSTTMVQVYVNMIDNKLKVTQQRKD-----NDIISSEYNTEMELHNYIK--D 159

QY	135	SEVNKLSSELKNAKEEINVLHGRLEKLNLVNNNNIENYVDSKVANLIFVNSLDGKCKC	194
DB	160	NLDNNIPSSLRVRAVIDSLHKKIQL-----ENATATQ-----DYCRSPC	201
QY	195	PSQEIQSRPVQHLYKDCSDYYAIGKRSETYRVTPQKSSFEVYCDMETMGGWTVL	254
DB	202	-----VASCNTPVSGRECEDIYRKGSETSEMYIIQDPPTTPYRVYCDMETMGGWTL	256
QY	255	QARLDGSTNFTRTWCQYKAGFGNLR-----EFWLGNDKIHLLTKSKEMILRID	303
DB	257	QNRQDGSVNFGRWDEYKRGFGNIAKSGGKKYCDTPGEYWLGNDKISQLTKIGPKVJIE	316
QY	304	LEDENGVELYALYDCFYVANEFLKYRLHVGNYNGTAGDALR--FNKHYN-----HDK	354
DB	317	MEDWNGDKVSALYGGFTIHNEGKYYQLSVSNYKGNAGNALREGASQLYGENRMTTHNGM	376
QY	355	FFTTPDKONDRY---PSGNCGLYYSSGWFFDACLSANLNGYY-----HQYRGVNR	403
DB	377	YFSTYCRDNDGWLTTDPRKQCKSKEDGGGWWYNRCHAANPNRGYYWGSTYSDMAKHGTDG	436
QY	404	GIFWGTWPVGSZAHPPGGYKSSFKEAKQMRP	434
DB	437	GIVWYXWKG-----SWY--SMKKMSNKKRP	459

```

RESULT 5
JC0596
fibrinogen-related protein HFREP-1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: JC0596
R:Yamamoto, T.; Gotoh, M.; Sasaki, H.; Terada, M.; Kitajima, M.; Hirohashi, S.
Biochem. Biophys. Res. Commun. 193, 681-697, 1993
A:Title: Molecular cloning and initial characterization of a novel fibrinogen-related
A:Reference number: JC0596; MUID:93290661; PMID:9390249
A:Accession: JC0596
A:Molecule type: mRNA
A:Residues: 1-312 <YAM>
A:Cross-references: GB:D14446; NID:G393314; PIDN:BAAC3336.1; PID:G393315
A:Experimental source: liver
C:Superfamily: fibrinogen beta/gamma homology
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-312/Product: fibrinogen-related protein HFREP-1 #status predicted <MAT>
F:80-305/Domain: fibrinogen beta/gamma homology <FE3>

```

```

Query Match      21.3%; Score 506.5; DB 2: Length 312;
Best Local Similarity 36.8%; Pred. No. 1e-27;
Matches 118; Conservative 38; Mismatches 108; Indels 57; Gaps 9

```

Qy	129	RVREIESEVNVKLSSELKNAKEEINVLHGRLEKINLVNMNIENIYVDSKVANSTFVANSLO	188
Db	36	QVRLLTRVKQQQVKIKQLLQ	73
Qy	189	GKSKCQPSQEIQSRPVQHLYKDCSDYYAIGKRSSEYRVTPQPKNSSFEVYCDMETNG	248
Db	74	-----LGSKRQ-----YADCSEIFNDGYKLSGFYKIKPLQSPAEFSVYCDXSD-G	117
Qy	249	GGWTVZQARLDGSTNFTRTWQCYKAGFGL--RRFPLGNDKIHLITKSKEMILRIDLE	305
Db	118	GGWTVIQRSDGSENFNRGWKDYENGFCNFVQKHGEYWGKNKJHFLTQEDYTLKIDLA	177
Qy	306	DFNGVELYALYDQFYVANEFLKYRLHVGNVNGTAGDALRFNKH-----YNHDLKFFTTP	359
Db	178	DFEKNRPAQYKQKFKVGDCKNFYELNIGEYSGTAGDSLACNFHPVEQVWASHQRMKFSTW	237
Qy	360	DKENDRYPSCNGCLYYSSGWWEFADCLSANLNGKYHHQKYRG-VRNGIFGWTGPGVSEHP	418
Db	238	DRDHNDY-EGNCAEEDQSGNWFNRCHSANLNGVYSGPYTAKTDNGIVWYTW-----	288
Qy	419	GGYKSSFKEAKQMIRPKHFKP	439
Db	289	HGWYYSJLKSVMKIRPNDPI	309

RESULT 6
EFGHJG3

fibrinogen gamma-B chain precursor [validated]; - human
 N;AlternateNames: coagulation factor I; fibrinogen gamma-55 chain
 C;Species: Homo sapiens (man)
 C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 08-Dec-2000
 C;Accession: A90494; A92448; A90453; A28203; B28203; I37390; A03126
 R;Pixon, M.W.; Chung, D.W.; Davie, E.W.
 Biochemistry 24, 2077-2086, 1985
 A;Title: Nucleotide sequence of the gene for the gamma chain of human fibrinogen.
 A;Reference number: A90494; MUID:85252774; PMID:2990550
 A;Accession: A90494
 A;Molecule type: DNA
 A;Residues: 1-113; Y', 115-453 <RIX>
 A;Cross-references: GB:M10014; GB:J00134; GB:J00135; GB:X00086; NID:gl82438; PIDN:AAB5
 R;Forrace Jr., A.C.; Cummings, D.E.; Comeau, C.M.; Kant, J.A.; Crabtree, G.R.
 J. Biol. Chem. 259, 12826-12830, 1984
 A;Title: Structure of the human gamma-fibrinogen gene. Alternate mRNA splicing near th
 A;Reference number: A92448; MUID:85010379; PMID:6092346
 A;Accession: A92448
 A;Molecule type: DNA
 A;Residues: 286-453 <FOR>
 R;Wolfenstein-Todel, C.; Mosesson, M.W.
 Biochemistry 20, 6146-6149, 1981
 A;Title: Carboxy-terminal amino acid sequence of a human fibrinogen gamma-chain varian
 A;Reference number: A90453; MUID:82068993; PMID:7306501
 A;Accession: A90453
 A;Molecule type: protein
 A;Residues: 411-434; Y', 436-440; Z', 442, Z', 444, B', 446-447, R', 449, ZBB', 453 <WOL>
 R;Francis, C.W.; Mueller, E.; Henschen, A.; Simpson, P.J.; Marder, V.J.
 Proc. Natl. Acad. Sci. U.S.A. 85, 3358-3362, 1988
 A;Title: Carboxyl-terminal amino acid sequences of two variant forms of the gamma-chai
 A;Reference number: A94194; MUID:88217900; PMID:3368448
 A;Accession: A28203
 A;Molecule type: protein
 A;Residues: 433-449 <FRA>
 A;Accession: B28203
 A;Molecule type: protein
 A;Residues: 433-453 <PR2>
 R;Marchetti, L.; Zanelli, T.; Malcovati, M.; Tenchini, M.L.
 DNA Seq. 1, 419-422, 1991
 A;Title: Polymorphism of the human gamma chain fibrinogen gene.
 A;Reference number: I37390; MUID:92119334; PMID:1685103
 A;Accession: I37390
 A;Status: translated from GB/EXBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 75-286 <RES>
 A;Cross-references: EMBL:X51473; NID:G31410; PIDN:CAA35837.1; P2D:G930064
 C;Comment: The two forms of gamma chain, A (see PIR:FGHUG) and B, arise by alternate s
 tron, which makes this chain different from the gamma-B chain at positions 434-437 an
 C;Comment: The gamma-B chain is present in about 10% of the fibrinogen molecules in pi
 C;Genetics:
 A;Gene: GDB:FGG
 A;Cross-references: GDB:119132; OMIM:134850
 A;Map position: 4q28-4q28
 A;Intons: 26/3; 41/3; 103/1; 134/2; 178/1; 222/3; 284/2; 377/1
 C;Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR:
 ins are contained in the core. Two three-chain coiled coils emerge from this core and
 from the distal domain nodes.
 C;Function:
 A;Description: fibrinogen cleaved by thrombin yields monomers that are polymerized int
 A;Pathway: blood coagulation
 C;Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
 C;Keywords: alternative splicing; blood coagulation; calcium; coiled coil; glycoprotei
 F;1-26/Domain: signal sequence #status predicted <SIG>
 F;27-453/Product: fibrinogen gamma-B chain #status experimental <MPT>
 F;176-415/Domain: fibrinogen beta/gamma homology <FBG>
 F;341-355/Domain: calcium binding #status predicted <CAB>
 F;400-422/Region: polymerization site, binding to the amino end of the alpha chain of
 F;34/Disulfide bonds: interchain (to gamma-35) #status predicted
 F;35/Disulfide bonds: interchain (to gamma-34) #status predicted
 F;45/Disulfide bonds: interchain (to beta-110) #status predicted

F;49/Disulfide bonds: interchain (to alpha-64) #status predicted	
F;78/Binding site: carbohydrate (Asn) (covalent) #status predicted	
F;161/Disulfide bonds: interchain (to beta-227) #status predicted	
F;165/Disulfide bonds: interchain (to alpha-180) #status predicted	
F;179-208,352-365/Disulfide bonds: #status predicted	
F;424/Cross-link: isopeptide (Gln) (interchain to Lys-432 N6-amino) #status predicted	
F;432/Cross-link: isopeptide (Lys) (interchain to Gln-424) #status predicted	
Query Match 20.4%; Score 484.5; DB 1; Length 453;	
Best Local Similarity 29.7%; Pred. No. 5;Se-26;	
Matches 141; Conservative 56; Mismatches 178; Indels 99; Gaps 18;	
QY 8 WLSSAVLATYGFVAVANNETEIEIKDERAKOVCPVLESRGKCEEAGECPVQVSLPPLTIQ 67	
DB 19 FLSSTCVA-----YVATRONCCILDERFGSYCPT-----TCGIADFLSTYQTX 61	
QY 68 LPKQFSRIEEVFKEVQNLKEIVNSLKKSCQDCKLQADONGDPGRNGLLLPSTGAPGEVGD 127	
DB 62 VDKDLOSLEDILHQVENKTSEVKQLIKAIQ-LTYNPDESSKPNMI-----DAAT 109	
QY 128 NRVRELESEVNLKLSSELKNAKEEINVHLHGRLEKLNLVMMNNIENYYVDSKVANLTFVNSL 187	
DB 110 LKSRKMLEEINKYEASILTHDSSIRYLQ-----ELYSNN-----QKIVNLKEKVAQL 157	
QY 188 DGKCKSCPSQEQIQSRPVQHLIYKDCSDYYAIGKSSSETYRVTPDPKNSSEFVYCDYETM 247	
DB 158 EAQCQE-PCKDTVQ---IHDITGKCQODIANKGAKGSCGYFIKPLKANQGFJYVCEIDGS 213	
QY 248 GGGWTVLQARLDGSTNFRRTQDYKAGFGNL----PREFWLGNDKIHLSJTKSKEM--ILR 301	
DB 214 GNGWTVFQKRLDGSVDKGMVIQYKEGFGHLSPITGTTEFWLGNKEIHLISTQSAIPYAJR 273	
QY 302 IDLEDFNGVELYALYDQFVAVNEFLKYRLHVGNV-NTAGCAL-RENKHVNHDLKFFT-- 357	
DB 274 VELEDWNGRTSTADYAMFKVGPEDAKYRLTYAYFAGSDAGDAFDGDFGDPSCKFFTSK 333	
QY 358 -----TPDKDNDRYPGNCGLYYSSSGWFDACLSANLNGKYH-----QKYRGVRNG 404	
DB 334 NGMCFSTWDNENCXKF-EGNCALQCGSGWYWKCHAGHLNGVYYOGGTYSKASTFNGYDNG 392	
QY 405 IFWGTNP-----GVSEA---HPGKYKSSFKEAKMMIRPKH 436	
DB 393 IINATWKTWYSYKKTMTKIIPNRLTIGEGQQHJGGAX-----QVRPEH 438	
RESULT 7	
FGHUG	
fibrinogen gamma-A chain precursor [validated] - human	
N;Alternate names: coagulation factor 1	
C;Species: Homo sapiens (man)	
C;Date: 24-Apr-1984 #sequence_revision 25-Feb-1985 #text_change 08-Dec-2000	
C;Accession: A90470; B90494; C94433; B93956; B92448; I37393; A40698; H54223; A03125; C37	
R;Chung, D.W.; Chan, W.Y.; Davie, E.W.	
Biochemistry 22, 3250-3256, 1983	
A;Title: Characterization of a complementary deoxyribonucleic acid coding for the gamma	
A;Reference number: A90470; MJID:83283434; PMID:6688357	
A;Accession: A90470	
A;Molecule type: mRNA	
A;Residues: 1-437 <CHU>	
R;Rixon, M.W.; Chung, D.W.; Davie, E.W.	
Biochemistry 24, 2077-2086, 1985	
A;Title: Nucleotide sequence of the gene for the gamma chain of human fibrinogen.	
A;Reference number: A90494; MJID:85252774; PMID:2990550	
A;Accession: B90494	
A;Molecule type: DNA	
A;Residues: 1-113,1',115-437 <RIX>	
A;Cross-references: GB:X02415; GB:M10014; NID:gl82438; PIDN:AAB59531.1; PID:gl82439	
R;Henschen, A.; Lottspeich, F.; Southan, C.; Topfer-Petersen, E.	
in Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe	
A;Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural	
A;Reference number: A94433	
A;Accession: C94433	
A;Molecule type: protein	

A;Residues: 27-437 <HEN>	
R;Kant, J.A.; Lord, S.T.; Crabtree, G.R.	
Proc. Natl. Acad. Sci. U.S.A. 80, 3953-3957, 1983	
A;Title: Partial mRNA sequences for human Aalpha, Bbeta, and gamma fibrinogen chains:	
A;Reference number: A93956; MJID:83247396; PMID:6575389	
A;Accession: B93956	
A;Molecule type: mRNA	
A;Residues: 276-437 <KAN>	
R;Fornace Jr., A.J.; Cummings, D.E.; Comeau, C.M.; Kant, J.A.; Crabtree, G.R.	
J. Biol. Chem. 259, 12826-12830, 1984	
A;Title: Structure of the human gamma-fibrinogen gene. Alternate mRNA splicing near th	
A;Reference number: A92448; MJID:85030379; PMID:6092346	
A;Accession: B92448	
A;Molecule type: DNA	
A;Residues: 286-437 <FOR>	
R;Imam, A.M.A.; Eaton, M.A.W.; Williamson, R.; Humphries, S.	
Nucleic Acids Res. 11, 7427-7434, 1983	
A;Title: Isolation and characterisation of cDNA clones for the Aalpha- and gamma-chain	
A;Reference number: I37393; MJID:84069777; PMID:6689067	
A;Accession: I37393	
A;Status: preliminary; translated from GB/EMBL/DDBJ	
A;Molecule type: mRNA	
A;Residues: 209-270 <RES>	
A;Cross-references: EMBL:X00086; NID:G31445; PIDN:CAA24944.1; PID:G577055	
R;Bertagnoli, M.E.; Beckerle, M.C.	
J. Cell Biol. 121, 1329-1342, 1993	
A;Title: Evidence for the selective association of a subpopulation of GPIIb-IIIa with	
A;Reference number: A40698; MJID:93286185; PMID:8509453	
A;Accession: A40698	
A;Molecule type: protein	
A;Residues: 27-33,'XX',36-41 <BER>	
A;Experimental source: thrombin-activated platelets	
A;Note: sequence extracted from NCBI backbone (NCBIP:133734)	
R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.	
Biochemistry 33, 1988-1993, 1994	
A;Title: Identification of proteins associated with apolipoprotein A-I-containing lipi	
A;Reference number: A54223; MJID:94162201; PMID:8117655	
A;Accession: H54223	
A;Molecule type: protein	
A;Residues: 27-33,'XX',36-41 <KUN>	
A;Note: identification of tryptic peptides from high-density lipoproteins	
R;Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.	
Ann. N. Y. Acad. Sci. 408, 28-43, 1983	
A;Title: Covalent structure of fibrinogen.	
A;Reference number: A90037; MJID:83254370; PMID:6575689	
A;Contents: annotation; review, disulfide bonds	
R;Doolittle, R.F.; Takagi, T.; Watt, K.; Bouma III, H.; Cottrell, B.A.; Cassman, K.G.;	
in Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S.; Ottesen, M.; Po	
A;Title: The structures of fibrinogen and fibrin.	
A;Reference number: A94437	
A;Contents: annotation; disulfide bonds	
R;Blomback, B.; Hessel, B.; Hogg, D.	
Thromb. Res. 8, 639-658, 1976	
A;Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.	
A;Reference number: A94309; MJID:76225080; PMID:936108	
A;Contents: annotation; disulfide bonds	
R;Hoeprich, P.D.; Doolittle, R.F.	
Biochemistry 22, 2049-2055, 1983	
A;Title: Dimeric half-molecules of human fibrinogen are joined through disulfide bonds	
A;Reference number: A90467; MJID:83231465; PMID:6860649	
A;Contents: annotation; quaternary structure, disulfide bonds	
R;Doolittle, R.F.	
Annu. Rev. Biochem. 53, 195-229, 1984	
A;Title: Fibrinogen and fibrin.	
A;Reference number: A90041; MJID:84305751; PMID:6383194	
A;Contents: annotation; review, EM structure, polymerization, ligands	
R;Horwitz, B.H.; Varadi, A.; Scheraga, H.A.	
Proc. Natl. Acad. Sci. U.S.A. 81, 5960-5984, 1984	
A;Title: Localization of a fibrin gamma-chain polymerization site within segment Thr-3	
A;Reference number: A94006; MJID:85014892; PMID:6592597	
A;Contents: annotation; polymerization region	
R;Kloczewiak, M.; Timmons, S.; Lukas, T.J.; Hawiger, J.	
Biochemistry 23, 1767-1774, 1984	

Db 146 VNLDRKVQLEANCOE-PCQDTVK-----IHDTGRCDQDVANKSAKESGLYFIRP-LKAKQ 200

QY 238 FEVYCDVETMGGWTVLQARLDGSTNFTRTWODYKAGFNL-----RREFWLGNDKIHL 291

Db 201 FLVYCEIDSGNGWTVFQKRLOGSLDPKKNWIOYKEGFGHISPTGTGNTFEWLGNFKIHL 260

QY 292 LTKSKEM--ILRIEDLEDFNGVELYALYDQFYVANEFLLKYRHVGMV-NGTAGDAJ----- 343

Db 261 ISTQSSIPVYLFICLEDWNGRTSTADYASFVKVTGENEYKRLVYAYFIGGDAGDAFDGVD 320

QY 344 ---RPNKHV-NHDJLKFTTTPDKDNDRYPGNGCGLYSSGWRFEACLSANLNGKYH----- 395

Db 321 GCDSSDKFTSHNGNQFSTWDSNDXY-DGNCAEQVGIGWMMNKKCHAGHLNGVYYQSGTY 379

QY 396 ---QKYRGVNRNGIFGWTPGVSEAHPGGYKS---SFKEAKMMIRP 434

Db 380 SKTSTPNGYDNG-IWATW-----KSRWYSKKXTTWKIIP 413

RESULT 11

FGHUB

fibrinogen beta chain precursor [validated] - human

N/Alternate names: coagulation factor I

N/Contains: fibrinopeptide B

C/Species: Homo sapiens (man)

C/Date: 24-Apr-1984 #sequence revision 31-Mar-1993 #text change 08-Dec-2003

C/Accession: B43568; A90469; B9C469; I37389; A94433; A9C437; A94309; G54223; A03121; B31

R/Chung, D.W.; Harris, J.E.; Davie, E.W.

Adv. Exp. Med. Biol. 281, 39-48, 1990

A/Title: Nucleotide sequences of the three genes coding for human fibrinogen.

A/Reference number: A43568; MUID:9134474C; PMID:2102623

A/Accession: B43568

A/Molecule type: DNA

A/Residues: 9-191,'P',193-491 <CHU>

R/Chung, D.W.; Que, B.G.; Rixon, M.W.; Mace Jr., M.; Davie, E.W.

Biochemistry 22, 3244-3250, 1983

A/Title: Characterization of complementary deoxyribonucleic acid and genomic deoxyribonu

A/Reference number: A90469; MUID:83283433; PMID:6688356

A/Accession: A90469

A/Molecule type: DNA

A/Residues: 1-38 <CH1>

A/Accession: B90469

A/Molecule type: mRNA

A/Residues: 9-191,'A',193-491 <CH2>

A/Cross-references: GB:J06129; NID:q182429; PIDN:AAA52429.1; PID:G182430

R/Huber, P.; Dalmon, C.; Courtois, G.; Laurent, M.; Assouline, Z.; Marguerie, G.

Nucleic Acids Res. 15, 1615-1625, 1987

A/Title: Characterization of the 5'-flanking region for the human fibrinogen beta gene.

A/Reference number: I37389; MUID:87146483; PMID:3029722

A/Accession: I37389

A/Status: translated from GB/EYEL/CDBJ

A/Molecule type: DNA

A/Residues: 1-38 <HUB>

A/Cross-references: EMBL:X05018; NID:q31400; PIDN:CAA28674.1; PID:g31401

R/Henschen, A.; Lottspeich, F.; Southan, C.; Topfner-Petersen, E.

in Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe

A/Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural v

A/Reference number: A94433

A/Contents: carbohydrate binding

A/Accession: A94433

A/Molecule type: protein

A/Residues: 31-137,'QS',140-144,'QF',147-491 <HEN>

R/Watt, K.W.K.; Takagi, T.; Doolittle, R.F.

Biochemistry 18, 68-76, 1979

A/Title: Amino acid sequence of the beta chain of human fibrinogen.

A/Reference number: A90437; MUID:79124640; PMID:420779

A/Accession: A90437

A/Molecule type: protein

A/Residues: 31-144,'QF',147-231,'D',233-330,'E',332-491 <WAT>

R/Blomback, B.; Hessel, B.; Hogg, D.

Thromb. Res. 8, 639-658, 1976

A/Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.

A/Reference number: A94309; MUID:76225080; PMID:936108

A/Contents: disulfide bonds

A/Accession: A94309

A/Molecule type: protein

A/Residues: 31-112,'E',114-137,'QS',140-144,'QF',147-148 <BLO>

R/Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.

Biochemistry 33, 1988-1993, 1994

A/Title: Identification of proteins associated with apolipoprotein A-I-containing lipo

A/Reference number: A54223; MUID:94162201; PMID:8117655

A/Accession: G54223

A/Molecule type: protein

A/Residues: 164-174 <KUN>

A/Note: Identification of tryptic peptides from high-density lipoproteins

R/Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.

Ann. N. Y. Acad. Sci. 408, 28-43, 1983

A/Title: Covalent structure of fibrinogen.

A/Reference number: A90037; MUID:83254370; PMID:6575689

A/Contents: annotation; review, disulfide bonds

R/Gardlund, B.; Hessel, B.; Marguerie, G.; Murano, G.; Blomback, B.

Eur. J. Biochem. 77, 595-610, 1977

A/Title: Primary structure of human fibrinogen. Characterization of disulfide-containi

A/Reference number: A91249; MUID:77245999; PMID:891553

A/Contents: annotation; disulfide bonds

R/Doolittle, R.F.; Takagi, T.; Watt, K.; Bouma II., H.; Cottrell, B.A.; Cassman, K.G.;

in Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S.; Ottesen, M., Fo

A/Title: The structures of fibrinogen and fibrin.

A/Reference number: A94437

A/Contents: annotation; disulfide bonds

R/Doolittle, R.F.

Annu. Rev. Biochem. 53, 195-229, 1984

A/Title: Fibrinogen and fibrin.

A/Reference number: A90041; MUID:84305751; PMID:6383194

A/Contents: annotation; review, EM structure, polymerization, ligands

R/Chung, D.W.; Rixon, M.W.; Que, B.G.; Davie, E.W.

Ann. N. Y. Acad. Sci. 408, 449-456, 1983

A/Title: Cloning of fibrinogen genes and their cDNA.

A/Reference number: A90038; MUID:83254384; PMID:6575700

A/Contents: annotation

R/Kirschbaum, N.E.; Budzynski, A.Z.

J. Biol. Chem. 265, 13669-13676, 1990

A/Title: A unique proteolytic fragment of human fibrinogen containing the Aalpha COOH-

A/Reference number: A37117; MUID:90337977; PMID:2143188

A/Contents: annotation; hementin cleavage site

A/Note: Hementin, a protease from Haemeteria ghilianii, the giant South American leec

C/Comment: The conversion of fibrinogen to fibrin is triggered by thrombin, which clea

ization sites responsible for the formation of the soft clot.

C/Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stabi

ger) and between alpha chains (weaker) of different monomers.

C/Comment: All fibrinogen chains are synthesized in the liver.

C/Genetics:

A/Gene: GDB:FGB

A/Cross-references: GDB:119130; OMIM:134830

A/Map position: 4q28-4q28

A/Introns: 38/3; 102/3; 164/1; 240/1; 278/1; 320/1; 415/2

C/Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR

ins are contained in the core. Two three-chain coiled coils emerge from this core and

from the distal domain nodes.

C/Function:

A/Description: fibrinogen cleaved by thrombin yields monomers that are polymerized int

A/Pathway: blood coagulation

C/Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disul

C/Keywords: blood coagulation; coiled coil; glycoprotein; liver; plasma; pyrogutamic

F;1-30/Domain: (or 4-30 or 15-30) signal sequence #status predicted <SIG>

F;31-491/Product: fibrinogen beta chain #status experimental <MAT>

F;31-44/Product: fibrinopeptide B #status experimental <API>

F;45-491/Product: fibrin beta chain #status experimental <FGB>

F;45-47/Region: polymerization site

F;99-228/Domain: fibrinogen disulfide ring homology <FDR>

F;238-487/Domain: fibrinogen beta/gamma homology <FBG>

F;31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experim

F;44-45/Cleavage site: Arg-Gly (thrombin) #status experimental

F;95/Disulfide bonds: interchain (to alpha-55) #status experimental

F;106/Disulfide bonds: interchain (to alpha-68) #status experimental

Db 241 LLSNLNGQYY--DYSGAPS--ZYWSYLFQDNDQ:P-----FAEMKLRNR 280

Search completed: November 5, 2003, 16:48:13
Job time : 28.0749 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 15:59:02 ; Search time 19.3392 Seconds
(without alignments)
1067.506 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MKLANWYLSAVLATYGF.....GYKSSFKKAGTMRPKHKP 439

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2378	100.0	439	1 FGL2_HUMAN	Q14314 homo sapien
2	1853.5	77.9	432	1 FGL2_MOUSE	P12804 mus musculu
3	525	22.1	496	1 AGP2_MOUSE	O35608 mus musculu
4	524.5	22.1	375	1 AGP2_BOVIN	Q77802 bos tauru
5	523	22.0	496	1 AGP2_PIG	Q9bdy7 sus scrofa
6	522	22.0	496	1 AGP2_HUMAN	Q15123 homo sapien
7	511.5	21.5	463	1 FIBB_CHICK	Q02020 gallus gall
8	506.5	21.3	312	1 FGL1_HUMAN	Q08930 homo sapien
9	486.5	20.5	498	1 AGP1_HUMAN	Q15389 homo sapien
10	485.5	20.4	493	1 ANL2_MOUSE	Q9xc45 mus musculu
11	484.5	20.4	453	1 FIBG_HUMAN	P02679 homo sapien
12	483.5	20.3	498	1 AGP1_MOUSE	O08538 mus musculu
13	480.5	20.2	493	1 ANL2_HUMAN	Q9ukug homo sapien
14	479.5	20.2	477	1 FIBB_PETMA	P02678 petromyzon
15	473	19.9	497	1 AGP1_RAT	O35460 rattus norv
16	471	19.8	468	1 FIBB_BOVIN	P02676 bos tauru
17	470.5	19.8	509	1 AGP4_BOVIN	Q9wvh6 mus musculu
18	467	19.6	481	1 AGP1_BOVIN	O18920 bos tauru
19	462	19.4	444	1 FIBG_BOVIN	P12799 bos tauru
20	461	19.4	491	1 FIBB_HUMAN	P02675 homo sapien
21	456	19.2	438	1 FIBG_XENLA	P17634 xenopus lae
22	454.5	19.1	479	1 FIBB_RAT	P14480 rattus norv
23	452	19.0	334	1 FCN1_MOUSE	O70165 mus musculu
24	451	19.0	866	1 FIBA_HUMAN	P02671 homo sapien
25	450	18.9	503	1 AGP4_HUMAN	Q9y264 homo sapien
26	446	18.8	282	1 FIBA_PARPA	P19477 parastichop
27	441	18.5	782	1 FIBA_RAT	P06399 rattus norv
28	438.5	18.4	335	1 FCN1_RAT	Q9wt58 rattus norv
29	428.5	18.0	432	1 FIBG_PETMA	P04115 petromyzon
30	428	18.0	319	1 FCN2_RAT	P57756 rattus norv
31	427.5	18.0	741	1 FIBA_CHICK	P14448 gallus gall
32	427	18.0	326	1 FCN1_HUMAN	O00632 homo sapien
33	421.5	17.7	306	1 FCN2_MOUSE	O70497 mus musculu

34	420	17.7	313	1 FCN2_HUMAN	Q15485 homo sapien
35	419	17.6	445	1 FIBG_RAT	P02680 rattus norv
36	414.5	17.4	299	1 FCN3_HUMAN	O75636 homo sapien
37	399	16.8	255	1 MFA4_HUMAN	P55083 homo sapien
38	397.5	16.7	4289	1 TENX_HUMAN	P22105 homo sapien
39	394.5	16.6	460	1 ANJ3_HUMAN	Q9y5ci homo sapien
40	388	16.3	2201	1 TENA_HUMAN	P24821 homo sapien
41	386.5	16.3	1746	1 TENA_PIG	Q29116 sus scrofa
42	386	16.2	1808	1 TENA_CHICK	P10039 gallus gall
43	374.5	15.7	406	1 ANL4_HUMAN	Q9by76 homo sapien
44	374.5	15.7	455	1 ANL3_MOUSE	Q9x182 mus musculu
45	369	15.5	410	1 ANL4_MOUSE	Q9z1p8 mus musculu

ALIGNMENTS

RESULT 1
FGL2_HUMAN
ID FGL2_HUMAN STANDARD; PRT; 439 AA.
AC Q14314;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinoleukin precursor (Fibrinogen-like protein 2) (p749).
GN FGL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=95369700; PubMed=7642106;
RA Ruegg C., Pytela R.;
RT "Sequence of a human transcript expressed in T-lymphocytes and
RT encoding a fibrinogen-like protein.";
RL Gene 160:257-262(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Yuvaraj S., Liu M., Marsden P., Levy G.;
RT "Cloning and characterization of Hfgl2: the human counterpart to the
RT mouse gene Fgl2.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLU-53.
RA Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA Poe C.L., Ozuna M., Yi Q., Nickerson D.A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schaefer C.F., Bhat N.K.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.C., McKernan K.J., Malek C.A., Gunaratne P.H.,
RA Villalón D.K., Muzry D.M., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahey C., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [5]
RP CHARACTERIZATION.
RX MEDLINE=98309432; PubMed=9647217;
RA Marazzi S., Blum S., Hartmann R., Gundersen D., Schreyer M.,
RA Argraves S., von Fliedner V., Pytela R., Ruegg C.;
RT "Characterization of human fibrocytes, a fibrinogen-like protein
secreted by T lymphocytes."
RL J. Immunol. 161:138-147(1998).
CC -!- FUNCTION: MAY PLAY A ROLE IN PHYSIOLOGIC LYMPHOCYTE FUNCTIONS AT
CC MUCOSAL SITES.
CC -!- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC
CC T-CELLS.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z36531; CAAB5298.1; -.
DR EMBL; AF104015; AAD10825.1; -.
DR EMBL; AF104014; AAD10825.1; JOINED.
DR EMBL; AF468953; AAL68855.1; -.
DR EMBL; BC033820; AAH33820.1; -.
DR PIR; I37391; I37391.
DR HSSP; P02671; 1FZD.
DR Genew; HGNC:3696; FGL2.
DR MIM; 605351; -.
DR GO; GO:0005576; C:extracellular; TAS.
DR GO; GO:0005973; C:fibrinogen beta chain; TAS.
DR GO; GO:0005974; C:fibrinogen gamma chain; TAS.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW T-cell; Glycoprotein; Signal; Polymorphism.
FT SIGNAL; 23 POTENTIAL.
FT CHAIN; 24 439 FIBROLEUKIN.
FT DOMAIN; 210 435 FIBRINOGEN C-TERMINAL.
FT DISULFID; 213 242 BY SIMILARITY.
FT DISULFID; 371 384 BY SIMILARITY.
FT CARBOHYD; 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD; 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD; 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD; 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD; 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT; 53 53 G -> E.
FT /FTID=VAR_013066.
FT SEQUENCE 439 AA; 50228 MW; DF34656288E49E68 CRC64;
Query Match 100.0%; Score 2378; DB 1; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.4e-153;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLANWYLSSAVLATYGFLLVANNETEEIKDERAKDVCPVLESRGKEAGECPYQVS 60
DB 1 MKLANWYLSSAVLATYGFLLVANNETEEIKDERAKDVCPVLESRGKEAGECPYQVS 60
QY 61 LPPLTIQLPKQFSRIEEVPEKVNKLKEIVNSLKSQCQCKLOADDNGDPGRNGLLPSTG 120
DB 61 LPPLTIQLPKQFSRIEEVPEKVNKLKEIVNSLKSQCQCKLOADDNGDPGRNGLLPSTG 120
QY 121 APGEVGNRVRELESEVKNLSSELKNAAKEEINVHGRLEKLNLVANNIENYVDSKVANL 180
DB 121 APGEVGNRVRELESEVKNLSSELKNAAKEEINVHGRLEKLNLVANNIENYVDSKVANL 180
QY 181 TFWVNSLDGKCKSPCEIQSRPVQHLIYKDCSDYYAIGKRSSETYRVTPDPKNSSFEV 240

Db 181 TFWVNSLDGKCKSPCEIQSRPVQHLIYKDCSDYYAIGKRSSETYRVTPDPKNSSFEV 240
QY 241 YCDMETMGGGWTVLQARLDGSGTNFTRTWQDYKAGFNLRRREFWLCNDKIHLTLTKSKEMIL 300
Db 241 YCDMETMGGGWTVLQARLDGSGTNFTRTWQDYKAGFNLRRREFWLCNDKIHLTLTKSKEMIL 300
QY 301 RIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRPNKHYNHDLKFFTTPD 360
Db 301 RIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRPNKHYNHDLKFFTTPD 360
QY 361 KNDRYPSGNCGLYSSGWWFDFACLSANLNGKYHQYRGVNRNGIFWGTWPGVSEAHPGG 420
Db 361 KNDRYPSGNCGLYSSGWWFDFACLSANLNGKYHQYRGVNRNGIFWGTWPGVSEAHPGG 420
QY 421 YKSSFEAKKMIRPKHKFP 439
Db 421 YKSSFEAKKMIRPKHKFP 439
RESULT 2
FGL2_MOUSE
ID FGL2_MOUSE STANDARD; PRT; 432 AA.
AC P12804;
DT 01-OCT-1989 (Rel. 12, Created);
DT 01-OCT-1989 (Rel. 12, Last sequence update);
DT 16-OCT-2001 (Rel. 40, Last annotation update);
DE Fibrocyte precursor (Fibrinogen-like protein 2) (Prothrombinase)
DE (Cytotoxic T-lymphocyte specific protein).
GN FGL2 OR FIBLP.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Cytotoxic T-cell;
RX MEDLINE=87175527; PubMed=3550794;
RA Koyama T., Hall J.R., Hasegawa W.G., Tonogawa S., Saito H.;
RT "Structure of a cytotoxic T-lymphocyte-specific gene shows a strong
RT homology to fibrinogen beta and gamma chains."
RL Proc. Natl. Acad. Sci. U.S.A. 84:1609-1613(1987).
RN 12;
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=BALB/CJ; TISSUE=Peritoneal macrophage;
RX MEDLINE=95333285; PubMed=7609073;
RA Parr R.L., Fung L., Reneker J., Myers-Mason N., Leibowitz J.L.,
RA Levy G.;
RT "Association of mouse fibrinogen-like protein with murine hepatitis
RT virus-induced prothrombinase activity."
RL J. Virol. 69:5033-5038(1995).
CC -!- FUNCTION: CONVERTS PROTHROMBIN TO THROMBIN.
CC -!- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC
CC T-CELLS.
CC -!- INDUCTION: IN MACROPHAGES, DURING INFECTION BY MOUSE HEPATITIS
CC VIRUS STRAIN 3 (MHV-3).
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16238; AAA37624.1; -.
DR EMBL; M15763; AAA37624.1; JOINED.
DR EMBL; S78773; AAB34823.1; -.
DR PIR; A27447; A27447.
DR HSSP; P02671; 1FZD.
DR MGD; MG1:103266; Fgl2.

```
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW T-cell; Cytolysis; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 FIBROLEUKIN.
FT DOMAIN 203 428 FIBRINOGEN C-TERMINAL.
FT DISULFID 206 235 BY SIMILARITY.
FT DISULFID 364 377 BY SIMILARITY.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 332 332 A -> G (IN REF. 2).
SQ SEQUENCE 432 AA; 48951 MW; 2B297F69CCB4A782 CRC64;

Query Match 77.9%; Score 1853.5; DB 1; Length 432;
Best Local Similarity 77.7%; Pred. No. 5.8e-118;
Matches 341; Conservative 42; Mismatches 49; Indels 7; Gaps 4;

QY 1 MKLANWYWLSSAVLATYGLFVANNETEIEIKDERAKDVCPVRLESRGKCEAGECPYQVS 60
Db 1 MRLPGWLWSSAVLAACR-AVEEHNLTGLEDSASAQAACPARLEGSGRC-EGSQCPFFQLT 58

QY 61 LPPLTIQLPKQFSRIEVEVKEVQNJKEIVNSLKKSCQDCKLQADDNGDPGRNGLLSPSTG 120
Db 59 LPTLTIQLPRQLGSMEEVLKEVRTLKEAVDSLKKSCQDCKLQADCHRDPPGGNG---GNG 114

QY 121 APGEVGNRVRELESEVKNLSSELKNAAKEEINVHGRLEKJNLVNMNNIENYVDSKVANL 180
Db 115 AE-TAEDSRVQELSESQVKNLSSELKNAAKQIQGLQGRLETLHLVNMNNIENYVDNKKVANL 173

QY 181 TFVNSLDGKCKSCPSQEQIQSRPVQHLYIKDCSDYYAIGKRSETYRVTPDPKNSSFEV 240
Db 174 TVVNSLDGKCKSCPSQEHYQSQPVQHLYIKDCSDHYVLGRRSSGAYRVTPDHRNSSFEV 233

QY 241 YCDYETMGGWTVLQARLDGSTNFTTRTWDYKAGFGNLRREFWLGNDKIHLTLTKSKEMIL 300
Db 234 YCDMETMGGWTVLQARLDGSTNFTREWKDYKAGFGNLEREFNLGNDKIHLTLTKSKEMIL 293

QY 301 RIDLEDENGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTPD 360
Db 294 RIDLEDENGLTLYALYDQFYVANEFLKYRLHIGNYNGTAGDALRFSRHHYNHDLRFFTPD 353

QY 361 KNDRYPSGNCGLLYSSGWNFDACLSANLNGKYYHQYKRGVRNGIFWGTWFGVSEAHPGG 420
Db 354 RNDRYPSGNCGLLYSSGWNFDSCLSANLNGKYYHQYKRGVRNGIFWGTWFGINQACPGG 413

QY 421 YKSSFKEAKVMIRPKFKP 439
Db 424 YKSSFQAKVMIRPKFKP 432

RESULT 3
AGP2_MOUSE STANDARD; PRT; 496 AA.
AC O356C8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Angiopoietin-2 precursor (ANG-2).
GN ANGPT2 OR AGPT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurograthi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=97349327; PubMed=9204896;
RA Maissonpierre P.C., Suri C., Jones P.F., Bartunkova S., Wiegand S.C.,
```

```
RA Radziejewski C., Compton D.L., McClain J., Aldrich T.H.,
RA Papadopoulos N., Daly T.C., Davis S., Sato T.N., Yancopoulos G.D.;
RT "Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo
RT angiogenesis.";
RJ Science 277:55-60(1997).
CC -!- FUNCTION: BINDS TO TIE2 RECEPTOR AND COUNTERACTS BLOOD VESSEL
CC MATURATION/STABILITY MEDIATED BY ANGIOPOIETIN-1. ITS FUNCTION MAY
CC BE CONTEXT-DEPENDENT. IN THE ABSENCE OF ANGIOGENIC INDUCERS, SUCH
CC AS VEGF, ANG2-MEDIATED LOOSENING OF CELL-MATRIX CONTACTS MAY
CC INDUCE ENDOTHELIAL CELL APOPTOSIS WITH CONSEQUENT VASCULAR
CC REGRESSION. IN CONCERT WITH VEGF, IT MAY FACILITATE ENDOTHELIAL
CC CELL MIGRATION AND PROLIFERATION, THUS SERVING AS A PERMISSIVE
CC ANGIOGENIC SIGNAL.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY AT SITES OF VASCULAR
CC REMODELING.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF004326; AAB63189.1; -.
DR HSSP; P02671; IFZD.
DR MGD; MGI:1202890; Agpt2.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SMC0186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Angiogenesis; Glycoprotein; Coiled coil; Signal.
FT SIGNAL 1 18. POTENTIAL.
FT CHAIN 19 496 ANGIOPOIETIN-2.
FT DOMAIN 159 256 COILED COIL (POTENTIAL).
FT DOMAIN 280 496 FIBRINOGEN C-TERMINAL.
FT DISULFID 284 313 BY SIMILARITY.
FT DISULFID 437 450 BY SIMILARITY.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 496 AA; 56616 MW; FA3021FE4E01C410 CRC64;

Query Match 22.1%; Score 525; DB 1; Length 496;
Best Local Similarity 32.9%; Pred. No. 3e-28;
Matches 127; Conservative 65; Mismatches 154; Indels 40; Gaps 9;

QY 61 LPPLTIQLPKQFSRIEVEVKEVQNJKEIVNSLKKSCQDCKLQADDNGDPGRNGLLSPSTG 120
Db 142 LTDVEAQVLNQTTRL-----ELQLLQHSISTNKLKQILDQTSINKLQNKNSFL----- 191

QY 121 APGEVGNRVRELESEVKNLSSELKNAAKEEINVHGRLE-----EKLNLVNMNNIENY 172
Db 192 -----EQKVLDMEGKHSEQLQSKYKEQKDELQVLVSKSSVIDELEKLVATVNN--SL 243

QY 173 VDSKVANLTFVNSLDGKCKSCPSQEQIQSRPVQHLYIKDCSDYYAIGKRSETYRVTPD 232
Db 244 LQXQQHDLMETVNSLLTMMSSPNSKSSVAIRKEEQTTFRDCAEIFKSGLTTSGIYLTFFP 303

QY 233 PKNSSFEVYCDMETMGGWTVLQARLDGSTNFTTRTWDYKAGFGNLRREFWLGNDKIHL 292
Db 304 NSTEETKAYCDMVGCGGWTVLQHRDGSVDQFQRTWKYEKFGFNPLGEYWLGNFEVSQ 363

QY 293 TKSKEMLRIDLEDENGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNH 352
Db 364 TGQHRVYLKIQKDWEGNEAHSLYDHFYLAGEESNYRIHLTGLGTAAKISSISQPGSD- 422

QY 353 LKFFTPDKDNDRYPSGNCGLLYSSGWNFDACLSANLNGKYYHQYKRGVR-NGIFWGTW 411
```



```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; AF233228; AAK14993.1; -.
DR HSSP; P02672; 1F2D.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00314; FIBRIN_AG_C DOMAIN; 1.
KW Angiogenesis; Glycoprotein; Coiled coil; Signal.
FT SIGNAL 1 18
FT CHAIN 19 496
FT DOMAIN 130 255 COILED COIL (POTENTIAL).
FT DOMAIN 280 496 FIBRINOGEN C-TERMINAL.
FT DISULFID 284 313 BY SIMILARITY.
FT DISULFID 437 450 BY SIMILARITY.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 496 AA; 56911 MW; 33B02BE224FE6B9D CRC64;

Query Match 22.0%; Score 523; DB 1; Length 496;
Best Local Similarity 34.9%; Pred. No. 4,le-28;
Matches 121; Conservative 49; Mismatches 120; Indels 57; Gaps 8;

QY 127 DNRVRELESEVWKLSS-----ELKNAKEEINVLHGRLEKJNLVNMN 168
Db 171 EKQILDQTSEINKLQDKNSFLEKKVLDMEKHIVQLQSIKEEKDQLQVLSKQNSIIEEL 230
QY 169 IENVYDSKVAN-----LTFVNSLQKCKSKCPSQEIQSRPVCHLIYK 211
Db 231 EKQLVTATVNSVLQKQCHDLMETVHNLLTMISTSNSAKHSLVAKEEQI-----IFR 282
QY 212 DCSYYAIGKRSSETYRTVPDKNSSFEVYCDMETMGGWTVLQARLDGSGTNFTRWQDY 271
Db 283 DCBAEAFKSLGTTSGTYTLTFPNSSTEETKAYCDMETGGGWTVIQRRDGSVDFTWKEY 342
QY 272 KAGFGNLRREFWLGNDKIHLTYKSKEMILRDLDEFNGVEYALYDQFYVANEFFKYRLH 331
Db 343 KMGFGSPSGEHWLGNFVSQVTNQKRYVLKIHLRDWEGNEAISLYEHFYJSSEEFNYRIH 402
QY 332 VGNVNGTAGDALRFNKHYNHDLKPFPTPDKCNDRYPSNGCGLYYSSGWWFDCLSANLNG 391
Db 403 LKGLGTAGKISSISQPGKD----FSTKDADNKKIC-KCSQMLTGGWTFDAGCPNSLNG 457
QY 392 KYVHQYRGVR-NGIFWGTVPGVSEAHPGGYKSSFKKAKMIRPKHF 437
Db 458 MYTFQRQNTNKFNGIKWYYWKG-----SGY--SLKA-TMMIRPADP 496

RESULT 6
AGP2_HUMAN
ID AGP2_HUMAN STANDARD; PRT; 496 AA.
AC O15123; Q9NRR7; Q9P2Y7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Angiopoietin-2 precursor (ANG-2).
GN ANGPT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
```

```
RX MEDLINE=97349327; PubMed=9204896;
RA Maisorpierre P.C., Suri C., Jones P.F., Bartunkova S., Wiegand S.J.,
RA Radziejewski C., Compton D., McClain J., Aldrich T.H.,
RA Papadopoulos N., Daly T.J., Davis S., Sato T.N., Yancopoulos G.D.;
RT "Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo
RT angiogenesis.";
RL Science 277:55-60(1997).
RM [2]
RX SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=99126459; PubMed=9927494;
RA Tanaka S., Mori M., Sakamoto Y., Makuuchi M., Sugimachi K.,
RA Wards J.P.;
RT "Biologic significance of angiopoietin-2 expression in human
RT hepatocellular carcinoma.";
RL J. Clin. Invest. 103:341-345(1999).
RM [3]
RX SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Umbilical vein endothelial cells;
RX MEDLINE=20309815; PubMed=10766762;
RA Kim J., Kim J.-H., Ryu Y.S., Jung S.H., Nah J.J., Koh G.Y.;
RT "Characterization and expression of a novel alternatively spliced
RT human angiopoietin-2.";
RL J. Biol. Chem. 275:18550-18556(2000).
CC -!- FUNCTION: BINDS TO TIE2 RECEPTOR AND COUNTERACTS BLOOD VESSEL
CC MATURATION/STABILITY MEDIATED BY ANGIOPOIETIN-1. ITS FUNCTION MAY
CC BE CONTEXT-DEPENDENT. IN THE ABSENCE OF ANGIOGENIC INDUCERS, SUCH
CC AS VEGF, ANG2-MEDIATED LOOSENING OF CELL-MATRIX CONTACTS MAY
CC INDUCE ENDOTHELIAL CELL APOPTOSIS WITH CONSEQUENT VASCULAR
CC REGRESSION. IN CONCERT WITH VEGF, IT MAY FACILITATE ENDOTHELIAL
CC CELL MIGRATION AND PROLIFERATION, THUS SERVING AS A PERMISSIVE
CC ANGIOGENIC SIGNAL.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O15123-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O15123-2; Sequence=VSP_001540;
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; AF004327; AAB63190.1; -.
DR EMBL; AB009865; BAA95590.2; -.
DR EMBL; AF187858; AAF76526.1; -.
DR HSSP; P02672; 1F2D.
DR Genew; HGNC:485; ANGPT2.
DR MIM; 601922; -.
DR GO; GO:0005625; C:extracellular space; TAS.
DR GO; GO:0005102; F:receptor binding activity; TAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00314; FIBRIN_AG_C DOMAIN; 1.
KW Angiogenesis; Glycoprotein; Coiled coil; Signal; Alternative splicing.
FT SIGNAL 1 16
FT CHAIN 17 496 ANGIOPOIETIN-2.
FT DOMAIN 130 256 COILED COIL (POTENTIAL).
FT DOMAIN 280 496 FIBRINOGEN C-TERMINAL.
FT DISULFID 284 313 BY SIMILARITY.
FT DISULFID 437 450 BY SIMILARITY.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```

FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 97 148 Missing (in isoform 2).
FT /FTid=VSP 0C1540.
FT CONFLICT 268 268 MISSING (IN REF. 2).
SQ SEQUENCE 496 AA; 56919 YN; 5642A58847A7385C CRC64;

Query Match 22.0%; Score 522; DB 1; Length 496;
Best Local Similarity 27.9%; Pred. No. 4.8e-28;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVQN 84
Db 26 MDSIGKKQYQVHGSCSYTFLLPEMDNCRSSSPYVSNVQADAPLEYDDSVQRLQVLEN 85
QY 85 LKE-----IVSLKKSCQCKLQADDNGDPGRNGLLPLSTGAPGEVGN---- 128
Db 86 IMENNTQWLKLENYIQDNMKKXVEIQNAVQN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRESEV-----NKLSELKNAKEEINVHGR---LEKL 161
Db 135 TAEQTRKLTQVDAQVNLQNTTRLEQLLEHSLSTNKEKQILDQTSINKLQDKNSFLEKK 194
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
Db 195 VLAMEDKHIIQLCSIKEEKDQLQVLVSKQNSIEELEKKIVATVNNSVLQKQKHLMET 254
QY 184 VNSLDGKCKSPSQEQIQSRFPVQHLIYKDCSDYYAIGKRSSSEYKRVTPDPKNSSEPEYCD 243
Db 255 VNNLLTMSTNSAKOPTVAKEEQISFQCAEVFKSGHTTNGYTLTFPNSTEEIKAYCD 314
QY 244 METMGSGWTVLQARLDGSTNFTRTWQYKAGFGNLRREFWLGNDKIHLTKSKEMILRID 303
Db 315 MEAGGGGWTIIQRREDGSDVDFQRTWKYKVGFGNPSGEYWLGNFVSQLTNQQRVYVKIH 374
QY 304 LEQFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFKPKHYNHDLKFFTTTPKXDN 363
Db 375 LKDWEQNEAYSLYEHFYLSSSEENYRIHLKGLTGAGKISSISQFGND---PSTKQGEN 430
QY 364 DRYPSGNGCYVSSGWFEACLSANLNGKYHHQYRGVR-NGIFAGTWPVGVSEAHFGYK 422
Db 431 DKCIC-KCSQVLTGGWFEACGSPSNLNGMYFQCRQNTNKFNGIKWYWKG-----SGY- 482
QY 423 SSFKEAKMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 7
FIBB_CHICK STANDARD; PRT: 463 AA.
ID FIBB_CHICK STANDARD; PRT: 463 AA.
AC Q02020;
DT 01-JUL-1993 (Rel. 26, Created;
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B]
DE (Fragment).
GN FGB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-13 AND 18-39.
RX MEDLINE=91182745; PubMed=2009266;
RA Weissbach J., Oddoux C., Procyk R., Griendinger G.;
RT "The beta chain of chicken fibrinogen contains an atypical thrombin
RT cleavage site.";
RL Biochemistry 30:3290-3294(1991).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
```

```

CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
CC CONVERTED INTO THE HARD CLOT BY FACTOR XIII WHICH CATALYZES THE
CC EPSILON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
CC (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
CC MONOMERS.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M58514; AAA48770.1; -.
CC PIR; A38463; A38463.
CC PDB; 1EI3; 10-MAY-00.
CC InterPro; IPR002181; Fibrinogen_C.
CC Pfam; PF00147; fibrinogen_C; 1.
CC SMART; SM00186; FBG; 1.
CC PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
CC Blood coagulation; Plasma; Platelet; Glycoprotein; Sulfation;
CC 3D-structure.
CC NON_TER : 1 FIBRINOPEPTIDE B.
CC PEPTIDE <2 17 FIBRINOGEN BETA CHAIN.
CC CHAIN 18 463 SUFEATION (BY SIMILARITY).
CC MOD_RES 5 5 CLEAVAGE (BY THROMBIN; RELEASE
CC SITE 17 18 FIBRINOPEPTIDE B).
CC FT DISULFID 69 69 INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
CC FT DISULFID 80 80 INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
CC FT DISULFID 84 84 INTERCHAIN (WITH GAMMA) (BY SIMILARITY).
CC FT DISULFID 197 197 INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
CC FT DISULFID 201 201 INTERCHAIN (WITH GAMMA) (BY SIMILARITY).
CC FT DISULFID 205 289 BY SIMILARITY.
CC FT DISULFID 215 244 BY SIMILARITY.
CC FT DISULFID 397 410 BY SIMILARITY.
CC FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 463 AA; 52678 MW; 2044CD49BA79EC7B CRC64;

Query Match 21.5%; Score 511.5; DB 1; Length 463;
Best Local Similarity 30.4%; Pred. No. 2.2e-27;
Matches 137; Conservative 71; Mismatches 154; Indels 89; Gaps 16;

QY 31 KDERAKDVCPVRLSRGKC---EEAGE-CPYQVSLPPLTQLPKQFSRIEEVFKEVQN- 84
Db 51 QDKQAMKKGPILYDPAGGCKHPDELGVLCPTGCE---LQTLLKQEKTVKPVRLDKDR 107
QY 85 -----LKEIVNSLKKSCQCKLQADDNGDPGRNGLLPLSTGAPGEVGNRVRELE 134
Db 108 VAKFSDTSTTMQYVNMIDNKLVKTKQKRD-----NDIILSEYNTMELHNYIK--D 159
QY 135 SEVKNLSSELKNAKEEINVHGRLEKLNLMNMNIENYVDSKVANLTFVNSLDGKCKSC 194
Db 160 NLQNNIPSSLRVLRAVIDSLHKKIQKL-----ENAIATQT-----DYCRSPC 201
QY 195 PSQEQIQSRPVCHLIYKDCSDYYAIGKRSSSEYKRVTPDPKNSSEPEYVCDMETGGWTVL 254
Db 202 -----VASCNIPVVGRECEDIYRKGETSEMYIIQDPDFTFYRVYCDMETDNGGWTLLI 256
QY 255 QARLDGSTNFTRTWQDYKAGFGNLR-----EFWLGNDKIHLTKSKEMILRID 303
Db 257 QNRQDGSVNFGRADWEYKRGFGNIAKSGGKKYCDTPGEYWLGNDKISQLTKIGPTKVLLIE 316
QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALR--FNKHYN-----HDLK 354
Db 317 MEDWNGDKVSALYGGFTTHNEGNKYLQSVSNYKGNAGNALMEGASQLYGENRTMTTHNGM 376
```

```

CY      355 FFTPDKDNDRY---PSGNGGLYSSGGWFDACLSANLNGKY-----HOKYRGVRN 403
          :|:|:|:|:| | | | | | | | | | | | | | | | | | | | | | | |
DB      377 YFSTYDRDNDGLTTPRKQCKSKEDGGGWYNRCHAANPNGRYYWGTYSDMAKISTDD 436

CY      404 GLEWGTWPGVSEAHFGGYKSSFKEAQMIRP 434
          |||!| | | | | | | | | | | | | | | | | | | | | | | |
DB      437 GIWMNRWKG-----SWY--SMKKMSMKIKP 459

```

```

RESULT 8
FGL1_HUMAN
ID FGL1_HUMAN STANDARD; ERT; 312 AA.
AC Q08830; Q36KW6; Q95QM6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinogen-like protein 1 precursor (Hepatocyte-derived fibrinogen-
DE related protein 1) (HfREP-1) (Hepassocin) (Hp-041).
GN FGL; OR HfREP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93290661; PubMed=8390249;
RA Yamamoto T., Gotoh M., Sasaki H., Terada M., Kitajima M.,
RA Hirohashi S.;
RT "Molecular cloning and initial characterization of a novel
RT fibrinogen-related gene, HfREP-1.";
RL Biochem. Biophys. Res. Commun. 193:681-687(1993).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=21363035; PubMed=11470158;
RA Hara H., Yoshimura H., Uchida S., Toyoda Y., Aoki M., Sakai Y.,
RA Morimoto S., Shickawa K.;
RT "Molecular cloning and functional expression analysis of a cDNA for
RT human hepassocin, a liver-specific protein with hepatocyte mitogenic
RT activity.";
RL Biochim. Biophys. Acta 1520:45-53(2001).
RN [3]
RP SEQUENCE FROM N.A.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14446; BAA03336.1; -.
DR EMBL; D87342; BAB7069C.1; -.
DR EMBL; BC007047; AAH07047.1; -.
DR PIR; CNG596; JN0596.
DR HSSP; PC2671; 1FZD.
DR Genew; HGNC:3695; FGL1.
DR XIM; 605776; -.
DR GO; GO:0005577; C:fibrinogen complex; TAS.
DR InterPro; IPR002191; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Signal.
FT SIGNAL 1 22
FT CHAIN 23 312
FT DOMAIN 78 305
FT DISULFID 26 26
FT DISULFID 83 112
FT DISULFID 248 261
FT CONFLICT 15 15 I -> T (IN REF. 3).
FT CONFLICT 69 69 N -> D (IN REF. 1).
FT CONFLICT 72 72 I -> V (IN REF. 1).
FT CONFLICT 105 105 P -> L (IN REF. 2).
SQ SEQUENCE 312 AA; 36391 MW; 26BC82124E6660C2 CRC64;

Query Match 21.3%; Score 506.5; DB 1; Length 312;
Best Local Similarity 36.8%; Pred. No. 2.9e-27;
Matches 119; Conservative 40; Mismatches 119; Indels 45; Gaps 10;

QY 131 RELESEVKNLSSELKNAKEEINVHGRLE----KLNLVMMNIENYVDSKVANLTFVNS 186
DQ 18 REI-SALEDCAEQOMRLRAQVRLLETRVKQQQVKIKQLLOENEVQFLDKGDENTVIDLG- 75
QY 187 LDGKCKSPSQEQIQSRFVQHLIYKCCSDYIAGKRSSETYRVTPDPKNSSFVEYCDMET 246
DQ 76 -----SKRQ-----YADCSEIFNDGYKLSGFYKIKPLQSPAEFVYCDMSD 116
QY 247 MGGGWTVLQAPLGDGSTNFTRTWQDYKAGFGNL---RRFWLGNDKIHLTTKSEVILRID 303
DQ 117 -GGGWTVIQRSDGSENFNRGWKDYENGFGNFVQKHGEYWLGNKLNHLFTTQEDYTLKID 175
QY 304 LEDFNGVELYALDQFYVANEFLKYRLHVGNNGTAGDALRFNKH-----YNHDLKPFT 357
DQ 176 LADFEKNSRYAQYKNFKVGDEKXNFYELNIGEYSGTAGDSLGNFHPFVQWQWASHQRMKPS 235
QY 358 TPEKNDNRYPSGNCGLYSSGWWFDACLSANLNGKYVHQYRG-VRNGIFWGTWPGVSEA 416
DQ 236 TWRDRHDNY-EGNCAEEDQSGWFWNFRCHSANLNGVYSGPYTAKTDNGIVWYTW----- 288
QY 417 HPGGYKSSFEKAKVMIRPCHFKP 439
DQ 289 --HGWNYSLKSVVMKIRPNDPIP 309

RESULT 9
AGP: HUMAN
ID AGP: HUMAN STANDARD; PRT; 498 AA.
AC Q15389;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Angiopoietin-1 precursor (ANG-1).
GN ANGPT1 OR KIAA0003.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT GLY-269 DEL.
RC TISSUE=Fetal lung;
RX MEDLINE=97134663; PubMed=8980223;
RA Davis S., Aldrich T.H., Jones P.F., Acheson A., Compton D.L., Jain V.,
RA Ryan T.E., Bruno J., Radziejewski C., Maissonpierre P.C.,
RA Yancopoulos G.D.;
RA "Isolation of angiopoietin-1, a ligand for the Tie2 receptor, by
RT secretion-trap expression cloning";
RL Cell 87:1161-1169(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RL DNA Res. 9:99-106(2002).
RN [3]
RP SEQUENCE OF 307-498 FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. I.
RT The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RT cell line KG-1.";
RL DNA Res. 1:27-35(1994).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT GLY-269 DEL.
RA Nakatsukasa M., Komai K., Shiozawa S.;
RT "Human angiopoietin-1 mRNA variant form.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A., AND VARIANT GLY-269 DEL.
RA Shan Z.X., Yu X.Y., Lin Q.Y., Fu Y.H., Tan H.H., Zheng M., Lin S.G.;
RT "Human angiopoietin-1 mRNA variant forms.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS AND ACTIVATES TIE2 RECEPTOR BY INDUCING ITS
CC TYROSINE PHOSPHORYLATION. IMPLICATED IN ENDOTHELIAL DEVELOPMENTAL
CC PROCESSES LATER AND DISTINCT FROM THAT OF VEGF. APPEARS TO PLAY A
CC CRUCIAL ROLE IN MEDIATING RECIPROCAL INTERACTIONS BETWEEN THE
CC ENDOTHELIUM AND SURROUNDING MATRIX AND MESENCHYME. MEDIATES BLOOD
CC VESSEL MATURATION/STABILITY. IT MAY PLAY AN IMPORTANT ROLE IN THE
CC HEART EARLY DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: GLYCOSYLATED.
CC -!- MISCELLANEOUS: IT MAY HAVE A POTENTIAL THERAPEUTIC UTILITY SINCE
CC IT CAN BE USED FOR SPECIFICALLY TARGETING TUMOR VASCULATURE OR FOR
CC PROMOTING ANGIOGENIC PROCESSES IN CERTAIN ORGANS SUCH AS AN
CC ISCHEMIC HEART.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U83508; AAB50557.1; -;
DR EMBL; D13628; BAA02793.2; ALT_INIT.
DR EMBL; AB084454; BAB91325.1; -;
DR EMBL; AY21504; AAM81745.1; -;
DR EMBL; AY124380; AAM92271.1; -;
DR HSSP; P02671; 1F2D.
DR Genew; HGNC:494; ANGPT1.
DR MIM; 601667; -;
DR GO; GO:0005102; F:receptor binding activity; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.

DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF0C147; fibrinogen_C; 1.
DR SMART; SMCC186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Angiogenesis; Glycoprotein; Coiled coil; Signal; Polymorphism.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 498 ANGIOPOIETIN-1.
FT DOMAIN 81 119 COILED COIL (POTENTIAL).
FT DOMAIN 153 261 COILED COIL (POTENTIAL).
FT DOMAIN 284 498 FIBRINOGEN C-TERMINAL.
FT DISULFID 286 315 BY SIMILARITY.
FT DISULFID 439 452 BY SIMILARITY.
FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 269 269 MISSING (IN CELL LINE T98G; MAY BE DUE TO
FT EXON SLIPPAGE).
FT /FTID=VAR 009940.
SQ SEQUENCE 498 AA; 57513 MW; 5D5FA63AEF6BE920 CRC64;

Query Match 20.5%; Score 486.5; DB 1; Length 496;
Best Local Similarity 32.3%; Pred. No. 1.2e-25;
Matches 131; Conservative 59; Mismatches 136; Indels 79; Gaps 16;

QY 61 LPPLTIQLPKQFSRIE-----EVPKEVQNLIKVINSLKSKQDCKLQADDNGDPGRN 112
DB 145 LTVETQVLNQTSRLEIQLLSENSJTYKLEKQLQQTNEILKHE-----KN 191

QY 113 GLLPLSTGAPGEVGENRVRELESEVNKLSELKNAKEEINVLHG-----RLEK-L 161
DB 192 SLL-----EHKILEME---GKKEELDTLKEEKENLQGLVTRQTYIIQELEKQL 237

QY 162 NLVNMNIENYVDSKVANLTFVNVSLDGKCSK---CPSQEQIQSRPVQHLJYKCCSDYY 217
DB 238 NRATTNN--SVLQKQQLLELMDTVHNLVNLCTKEGVLLKGGKREEKP----FRDCADV 290

QY 218 AIGKRSETYRV---TPDPKNSSEYVCDYETMGGGWTVLQARLDGSTNFTRTWQDYKA 273
DB 291 QAGFNKSGIYTIYINNMPEPK----KVFNCYDVNGGGWTVIQHREDGSLDFCRGWKEYKM 346

QY 274 GFGNLRREFWLGNDKIHLTKSKEMILRIDJEDFNGVELYALYDQFYVANEFKYLRLHVG 333
DB 347 GFGNPSGEYWLGNESFAITSQRQYMLRIELMDWEGNRAYSQYDRFHIGNEKQNYRLYJK 406

QY 334 NYNGTAGDALRENNKHYNHDLKFFTTTPDKNDRYPSGNCGLYYSSGWFEDACLSANLNGKY 393
DB 407 GHTGTAGKQSSLLIH-GAD---FSTKDAENDNCMC-KCALNLTGGWFDACGPNJNGMF 461

QY 394 YHQ-KYRGVRNGIFWGTWPGVSEAHPGYKSSFKKAKMMIRPKHF 437
DB 462 YTAGQNHGKLG-KWHYFKGPS-----YSLRSTTMMIRPLDF 498

RESULT 10
ANL2_MOUSE
ID ANL2_MOUSE STANDARD; PRT; 493 AA.
AC Q9R045;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Angiopoietin-related protein 2 precursor (Angiopoietin-like 2).
GN ANGPTL2 OR ARP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=99403103; PubMed=10473614;
RA Kim I., Moon S.-O., Koh K.N., Kim H., Uhm C.-S., Kwak H.J., Kim N.-G.,

```
RA Koh G.Y.;
RT "Molecular cloning, expression, and characterization of angiopoietin-
RT related protein. angiopoietin-related protein induces endothelial cell
RT sprouting.";
RL J. Biol. Chem. 274:26523-26528(1999);
CC -!- FUNCTION: INDUCES SPROUTING IN ENDOTHELIAL CELLS THROUGH AN
CC AUTOCRINE AND PARACRINE ACTION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN HEART, TONGUE, LUNG AND
CC SKELETAL MUSCLE. ALSO FOUND IN LOWER LEVELS IN KIDNEY, EPIDIDYMUS
CC AND TESTIS.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF125176; AAD55358.1; -
DR HSSP; P02671; 1FZD.
DR YGD; MGI:134702; Angptl2.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Signal; Coiled coil; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 493 ANGIOPOIETIN-RELATED PROTEIN 2.
FT DOMAIN 77 115 COILED COIL (POTENTIAL).
FT DOMAIN 152 202 COILED COIL (POTENTIAL).
FT DOMAIN 438 450 FIBRINOGEN C-TERMINAL.
FT DISULFID 278 307 BY SIMILARITY.
FT DISULFID 430 443 BY SIMILARITY.
FT CARBOHYD 164 164 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 493 AA; 57118 MW; 22885ABEF0746BF2 CRC64;

Query Match 20.4%; Score 485.5; DB 1; Length 493;
Best Local Similarity 30.8%; Pred. No. 1.4e-25;
Matches 117; Conservative 62; Mismatches 128; Indels 73; Gaps 12;

QY 66 LQPKGFSRIEEVFKEV---QNLKEIVNSLKKSCQDKLOADMGDPGRNGLLPSTGA 121
DB 170 LQLASKYKDLHKFQHLAMLAHQSEVIAQLEEHQCRV-----PAARFMPQPPFAA 220
QY 122 PGEVGD---NRVRELESEVKNLSSELKNAKEEINVLHGRLEKLNLVNKNNIENYVDSKV 177
DB 221 PPRVYQPTYNRI-----INQISTNEIQSDQNKVLPPSLPTMPAL----- 261
QY 178 ANLTFVNSLDGKCKSPSQEQICSRPVQHLIYKDCSDYVAIGKRSSSEYRVTDPDKSS 237
DB 262 -----TSLPSSTDKPSGP-----WRDCLQALEDGHTSTSSIYLVKPEINRL 302
QY 238 FEVYCDMETMGGWTVLQARLDGSTNFTRTWCQDYKAGFGNLRREFWLGNDKIHLTKSKE 297
DB 303 XQVWCQDPHDPGGWTVICRRLDGSVNFRRNWTYKQFGNIDGHWLGLNLYWLTNGCN 362
QY 298 MLLRIDLEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDAAFRNKNYHDLKPEFT 357
DB 363 YKLLVTMEDWSGRKKVFAEYASFRLEPSEYKRLRGYHGNAGDSFTW-----HNGKQFT 417
QY 358 TPDKNDNRYPSGNCGLYSSGWFDACLSAKNGKYV---HQKYRGVNRNGIFWGTPGVGS 414
DB 418 TLDKRDHDVY-TGNCAHYQKGGWYNACAHSNLNGWYRGGHVRSR-YQDGVYWAEPFG-- 473
QY 415 EAHPGGYKSSFKKAKMVIKP 434
DB 474 ----GSY--SLKKVVMVIRP 487
```

```
RESULT 11
FIBG_HUMAN
ID FIBG_HUMAN STANDARD; PRT; 453 AA.
AC P02679; P04469; P04470; Q96A14; Q96KJ3;
DT 21-JUL-1986 (Rel. 01, Created)
DT 29-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinogen gamma chain precursor (PRO2061).
GN FBG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS GAMMA-A AND GAMMA-B).
RX MEDLINE=85252774; PubMed=2990550;
RA Rixon M.W., Chung D.W., Davie E.W.;
RT "Nucleotide sequence of the gene for the gamma chain of human
RT fibrinogen.";
RL Biochemistry 24:2077-2086(1985).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=83283434; PubMed=6688357;
RA Chung D.W., Chan W.-Y., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT the gamma chain of human fibrinogen.";
RL Biochemistry 22:3250-3256(1983).
RN [3]
SEQUENCE FROM N.A. (ISOFORMS GAMMA-A AND GAMMA-B).
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A. (ISOFORM GAMMA-A).
RX TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Bi J., Zhang Y., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 33 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
SEQUENCE FROM N.A. (ISOFORM GAMMA-A).
RX TISSUE=Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guraratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska J., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
SEQUENCE OF 27-437.
RA Henschen A., Lottspeich F., Southan C., Topfer-Petersen E.;
RT "Human fibrinogen: sequence, sulfur bridges, glycosylation and some
RT structural variants.";
RJ (In) Peeters H. (eds.);
RJ Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,
RJ Pergamon Press, Oxford (1980).
RN [7]
```



```
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Boraudo M.F., Casavant T.B., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.C., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.C.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: INDUCES SPROUTING IN ENDOTHELIAL CELLS THROUGH AN
CC AUTOCRINE AND PARACRINE ACTION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN HEART, SMALL INTESTINE,
CC SPLEEN AND STOMACH. ALSO FOUND IN LOWER LEVELS IN COLON, OVARY,
CC ADRENAL GLAND, SKELETAL MUSCLE AND IN PROSTATE.
CC -!- PTM: N-GLYCOSYLATED.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF1251175; AAD55357.1; -
CC EMBL; BC012368; AAH12368.1; -
CC HSSP; P02671; 1FZD.
CC
CC Genew; HGNC:490; ANGPT12.
CC
CC MIM; 605001; -
CC
CC GO; GO:0005615; C:extracellular space; TAS.
CC
CC GO; GO:0005102; F:receptor binding activity; TAS.
CC
CC GO; GO:0007275; P:development; TAS.
CC
CC InterPro; IPR002181; Fibrinogen_C.
CC
CC Pfam; PF00147; fibrinogen_C; -
CC
CC SMART; SM00186; FBG; -
CC
CC PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
CC
CC Signal; Coiled coil; Glycoprotein.
CC
CC SIGNAL 1 22 POTENTIAL.
CC CHAIN 23 493 ANGIOPOIETIN-RELATED PROTEIN 2.
CC DOMAIN 76 115 COILED COIL (POTENTIAL).
CC DOMAIN 152 236 COILED COIL (POTENTIAL).
CC DOMAIN 438 450 FIBRINOGEN C-TERMINAL.
CC DOMAIN 278 307 BY SIMILARITY.
CC DISULFID 430 443 BY SIMILARITY.
CC DISULFID 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 493 AA; 57104 MW; CF2ADECESD185CA CRC64;
CC SEQUENCE
CC
CC Query Match 20.2%; Score 480.5; DB 1; Length 493;
CC Best Local Similarity 30.8%; Pred. No. 3e-25;
CC Matches 117; Conservative 62; Mismatches 128; Indels 73; Gaps 12;
CC
CC QY 66 IQLPKQFSRIEVEFKV---QNCKEIVNLSKKSCQCKLQADDNGDPGRNGLLPSTGA 121
CC :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
CC Db 170 LQLASKYKOLEHKYQHLATLAHQSEIIAQBEHCQV-----PSARVPQPPAA 220
CC
CC QY 122 PGEVGD---NRVRESEWNKLSSELKNKAEINVLGRLEKLNLVKNMNIENVDSKV 177
```

```
Db 221 PPRVYQPTYNRI-----INQISTNEIQSDQNLKVLPPPLPTMPTL----- 261
QY 178 ANLTFVNSLDGKSKCPSQEQIQSRPVOHLYKDCSDYVAIGKSSSEYRVTPDPKNS 237
Db 262 -----TSLPSSTDKPSGP-----WRDCLQALEGDHDTSSIYLVKPTNRL 302
QY 238 FEVYCDMETMGGWTVLQARLDGSTNFTPTWQDYKAGFNLRRREFWLGNDKIHLTKSKE 297
Db 303 MQVWCDQRHDPGGWTVIQRRLDGSVNFPRWETYSKQFGNIDGEYWLGLNIYWLTNQGN 362
QY 298 MILRIDLEDENGVELYALYDQFYVANEFLKYRLHVGNNGTAGDALRPNKHYNHDLKFF 357
Db 363 YKLLVTMEDWSGRKVFAYEASFRLPESEYKLRGGRYHGNAGDSFTW-----HNGKQFT 417
QY 358 TPKDNDRIYPSGKCLYSSGWFDACLSANLNGKYY---HQKYRGVRNCIFWGTWPGVS 414
Db 418 TCDRDHDVY-TGNCAHYQKGGWYNACAHNSLNGVWYRGGHYRSR-YQDGVYMAEFRG-- 473
QY 415 EAHPPGGYKSSPFKEAKMIRP 434
Db 474 ----GSY--SLKKVVMIRP 487

RESULT 14
FIBB_PETVA
ID FIBB_PETVA STANDARD; PRT; 477 AA.
AC P02678;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragments).
OS Petromyzon marinus (Sea lamprey).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
CC Petromyzontiformes; Petromyzontidae; Petromyzon.
CC NCBI_TaxID=7757;
RN [1];
RP SEQUENCE OF 1-36.
RX MEDLINE=77065679; PubMed=999898;
RA Cottrell B.A., Doolittle R.F.;
RT "Amino acid sequences of lamprey fibrinopeptides A and B and
RT characterizations of the junctions split by lamprey and mammalian
RT thrombins.";
RL Biochim. Biophys. Acta 453:426-438(1976).
RN [2];
RP SEQUENCE OF 37-477 FROM N.A.
RX MEDLINE=87076582; PubMed=3790537;
RA Bohonus V.L., Doolittle R.F., Pontes M., Strong D.D.;
RT "Complementary DNA sequence of lamprey fibrinogen beta chain.";
RL Biochemistry 25:6512-6516(1986).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NON-IDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
CC CONVERTED INTO THE HARD CLOT BY FACTOR XIIIa WHICH CATALYZES THE
CC EPSILON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
CC (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
CC MONOMERS.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
```

DR EMBL; M14773; AAA49261.1; --
DR PIR; A25052; A25052.
DR PDB; 1LWU; 23-AUG-02.
DR PCB; 1N73; 07-CAN-03.
DR InterPro; IPR002181; Fibrinogen_C.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Blood coagulation; Plasma; Sulfation; Glycoprotein; 3D-structure.
FT PEPTIDE 1 36 FIBRINOPEPTIDE B.
FT NCN CONS 35 37
FT CHAIN <37 477 FIBRINOGEN BETA CHAIN.
FT MCD RES 13 13 SULFATION.
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .).
FT DISULFID 84 84 INTERCHAIN (WITH ALPHA CHAIN);
FT (BY SIMILARITY).
FT DISULFID 95 95 INTERCHAIN (WITH ALPHA CHAIN);
FT (BY SIMILARITY).
FT DISULFID 99 99 INTERCHAIN (WITH GAMMA CHAIN);
FT (BY SIMILARITY).
FT DISULFID 212 212 INTERCHAIN (WITH ALPHA CHAIN);
FT (BY SIMILARITY).
FT DISULFID 216 216 INTERCHAIN (WITH GAMMA CHAIN);
FT (BY SIMILARITY).
FT DISULFID 220 304 BY SIMILARITY.
FT DISULFID 230 259 BY SIMILARITY.
FT DISULFID 412 425 BY SIMILARITY.
SQ SEQUENCE 477 AA; 54270 MW; B8A95E7E32D09D18 CRC64;

Query Match 20.2%; Score 479.5; DB 1; Length 477;
Best Local Similarity 29.2%; Pred. No. 3.4e-25;
Matches 131; Conservative 66; Mismatches 138; Indels 113; Gaps 15;

QY 29 EIKDERAKDVCVRLESRGKCEEAGECPYQVSLPPTLQLPKQFSRIEEVFKEVQNLEKI 88
Db 100 ELREELLKQDPVR-----YKISMLK-----QNLTYF 126

QY 99 VNSLKSCQCKLQADDNGDPGRNGLLPSTGAPGEVGNRVRELESEWNKLSSELKAK 148
Db 127 INSFDR-----MASDSNTLKQ-----VQLRRRLNSRSTHYNAQ 162

QY 149 EEINVLHGLEKLNLMNNTIENVYDSKVANLTFVNSLDGKCK-----CPSQEQIQSRP 204
Db 163 KEI---ENYKEVKI---RIESTVAGSLRSMKSVLEHLRAKQMPMEEAIAKTQKELCSAP 215

QY 205 -----VQHLYKDCSDYYAIGKRSSERYAVTPDPKNSSFEVYCDMETMGGGWTVLQARL 256
Db 216 CTVNCRVPVVSGMHCEDIYRNGGRTSEAYIOPDLFSEPYKVFCDMESHGGGWTVVQNRV 275

QY 259 DGSTNFTRTQDYKAGFGNLR-----REFWLGNDKIHLLTKSKEMILRIDJEDF 307
Db 276 DGSSNFARDNTYKAIEFGNIAFGNGKSGICNIPGEYWLGTKTVHQLTKQHTQQVLFDMSDW 335

QY 308 NGVELYALVDQYVANEFLKYLRLHVGNYNGTAGDAL-----RFNKHYN-HDLKFEFT 358
Db 336 EGSSVYAQYASFRPENEAAQGYRLWVEDYSGNAGNALLEGATQLXGDNRTMTIHHGMQFST 395

QY 359 PDKNDRY-----PSGNCGLYSSGWFFDACL SANLNGKY-----HOKYRGVNRNGFW 407
Db 396 FDRDNNMNPGDPTKCHSREDAGGWYWRCHAANFNRYWGGYVYKEADYGTDCGVVW 455

QY 408 GTWPGVSEAHFPGYKSSFKEAKVMIRPK 435
Db 456 MNWKG-----SWY--SMRQAMKLRPK 475

RESULT 15
AGPI_RAT
ID AGPI_RAT STANDARD; PRT; 497 AA;
AC O35460; Q8K4Q4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Angiopoietin-1 precursor (ANG-1).

GN ANGPT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Placenta;
RX MEDLINE=22346496; PubMed=12458684;
RA Iizasa H., Bae S.H., Asashima T., Kitano T., Matsunaga N.,
RA Terasaki T., Kang Y.S., Nakashima E.;
RT "Augmented expression of the tight junction protein occludin in brain
endothelial cell line FR-hb5 by rat angiopoietin-1 expressed in
RT baculovirus-infected sf plus insect cells.";
RL Pharm. Res. 19:1757-1760(2002).
RN {2}.
RP SEQUENCE OF 91-200 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Placenta;
RX MEDLINE=98451564; PubMed=9776732;
RA Mandriota S.J., Pepper M.S.;
RT "Regulation of angiopoietin-2 mRNA levels in bovine microvascular
endothelial cells by cytokines and hypoxia.";
RL Circ. Res. 83:852-859(1998).
CC -!- FUNCTION: BINDS AND ACTIVATES TIE2 RECEPTOR BY INDUCING ITS
CC TYROSINE PHOSPHORYLATION. IMPLICATED IN ENDOTHELIAL DEVELOPMENTAL
CC PROCESSES LATER AND DISTINCT FROM THAT OF VEGF. APPEARS TO PLAY A
CC CRUCIAL ROLE IN MEDIATING RECIPROCAL INTERACTIONS BETWEEN THE
CC ENDOTHELIUM AND SURROUNDING MATRIX AND MESENCHYME. MEDIATES BLOOD
CC VESSEL MATURATION/STABILITY. IT MAY PLAY AN IMPORTANT ROLE IN THE
CC HEART EARLY DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB080023; BAC10290.1; --
CC EMBL; AF030376; AAC78246.1; --
CC InterPro; IPR002181; Fibrinogen_C.
CC Pfam; PF00147; fibrinogen_C; 1.
CC SMART; SM00186; FBG; 1.
CC PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Angiogenesis; Glycoprotein; Coiled coil; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 497 ANGIOPOIETIN-1.
FT DOMAIN 81 119 COILED COIL (POTENTIAL).
FT DOMAIN 153 261 COILED COIL (POTENTIAL).
FT DOMAIN 283 497 FIBRINOGEN C-TERMINAL.
FT DISULFID 285 314 BY SIMILARITY.
FT DISULFID 438 451 BY SIMILARITY.
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 98 98 Q -> E (IN REF. 2).
FT CONFLICT 172 172 E -> K (IN REF. 2).
FT CONFLICT 189 189 E -> K (IN REF. 2).
SQ SEQUENCE 497 AA; 57461 MW; 08E66AE8EFD868AE CRC64;

Query Match 19.9%; Score 473; DB 1; Length 497;
Best Local Similarity 30.1%; Pred. No. 9.7e-25;
Matches 125; Conservative 69; Mismatches 121; Indels 100; Gaps 16;

QY 61 LPPLTIQLPKQFSRIEEVFKEVQNLEKIVNS-----LKKSCQCKLQADDNGDPGRNG 113
Db 145 LTDVETQVLNQTSRL-----EIQLLENSLS:YELEKQLLQQTNEILKIQE-----KNS 192

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 16:37:23 ; Search time 58.0176 Seconds
(without alignments)
1952.598 Million cell updates/sec

Title: US-09-902-563-2
Perfect score: 2378
Sequence: 1 MKLANWYWLSSAVLATYGFJ.....QYKSSFKERKMMIRPKHFKP 439

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs. 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_undefined:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2162.5	90.9	442	6	Q8MIP7	Q8mip7 sus scrofa
2	1647	69.3	357	11	Q9EPT7	Q9ept7 rattus norv
3	750	31.5	148	4	Q8WWE4	Q8wwe4 homo sapien
4	618.5	26.0	652	5	Q9NDQ1	Q9ndq1 ciona intes
5	593.5	25.0	220	5	Q8T8A2	Q8t8a2 ciona savig
6	526	22.1	496	11	Q9D2D2	Q9d2d2 mus musculu
7	524	22.0	407	13	Q9DER1	Q9der1 gallus gall
8	524	22.0	441	13	Q9DER0	Q9der0 gallus gall
9	524	22.0	493	13	Q9DER2	Q9der2 gallus gall
10	523	22.0	407	13	Q9PU54	Q9pu54 gallus gall
11	518.5	21.8	314	11	Q8VC25	Q8vc25 mus musculu
12	506	21.3	312	4	Q8NG32	Q8ng32 homo sapien
13	505	21.2	491	4	Q95841	Q95841 homo sapien
14	496	20.9	244	4	Q9HBP3	Q9hbp3 homo sapien
15	491.5	20.7	513	13	Q90Z19	Q90z19 brachydanio
16	491	20.6	337	11	Q8R1Q3	Q8r1q3 mus musculu

17	489	20.6	346	4	O43827	O43827 homo sapien
18	485.5	20.4	292	5	Q9U8W8	Q9u8w8 tachypleus
19	485.5	20.4	493	11	Q8BM09	Q8bm09 mus musculu
20	483.5	20.3	498	11	Q8C2K6	Q8c2k6 mus musculu
21	482.5	20.3	488	13	Q91589	Q91589 xenopus lae
22	481.5	20.2	493	11	Q9JC03	Q9jj03 rattus norv
23	476	20.0	308	5	Q9U8W6	Q9u8w6 tachypleus
24	475	20.0	435	13	Q93568	Q93568 gallus gall
25	474.5	20.0	498	6	Q9BDY8	Q9bdy8 sus scrofa
26	473	19.9	489	13	Q9CZ18	Q90z18 brachydanio
27	471	19.8	316	5	Q9J8W7	Q9u8w7 tachypleus
28	460.5	19.4	461	4	Q8N539	Q8n539 homo sapien
29	456.5	19.2	324	5	Q95P99	Q95p99 halocynthia
30	456	19.2	341	5	Q966W1	Q966w1 halocynthia
31	456	19.2	481	11	Q8K0E8	Q8k0e8 mus musculu
32	455.5	19.2	236	4	Q8N2J9	Q8n2j9 homo sapien
33	454.5	19.1	431	6	Q95LU3	Q95lu3 macaca fasc
34	454	19.1	356	5	Q95P98	Q95p98 halocynthia
35	448	18.8	436	11	Q8VCM7	Q8vcm7 mus musculu
36	447.5	18.8	457	11	Q8R0Z6	Q8r0z6 mus musculu
37	445.5	18.7	269	11	Q8BJE7	Q8bje7 mus musculu
38	443.5	18.7	269	11	Q8RL54	Q8rl54 mus musculu
39	439.5	18.5	470	4	Q9BZZ0	Q9bzz0 homo sapien
40	439.5	18.5	470	4	Q8NI99	Q8ni99 homo sapien
41	437	18.4	337	6	Q9BE00	Q9be00 macaca fasc
42	435.5	18.3	235	6	Q28763	Q28763 papio cynoc
43	434.5	18.3	326	6	Q29042	Q29042 sus scrofa
44	434	18.3	200	11	Q8BMV1	Q8bmv1 mus musculu
45	433.5	18.2	324	5	Q95PA0	Q95pa0 halocynthia

ALIGNMENTS

RESULT 1

Q8MIP7
ID Q8MIP7 PRELIMINARY; PRT; 442 AA.
AC Q8MIP7;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DI 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Fibrinogen-like protein 2.
GN FGL2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Ghanekar A., Liu H., Grant D.R., Levy G.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY112657; AAM52324.1; -
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 442 AA; 50579 MW; 67800D67AEDFF899 CRC64;

Query Match 90.9%; Score 2162.5; DB 6; Length 442;
Best Local Similarity 89.1%; Pred. No. 8.5e-137;
Matches 392; Conservative 27; Mismatches 20; Indels 1; Gaps 1;

QY	1	MKLANWYWLSSAVLATYGFVANNETEETIKDERAKDVCPVRLESRGKCEEGECYPQVS	60
DB	3	JKLSNWCWLSSAVLAAYGFLVANNETEETIKDEAAKADCPVRLESRGKCEEGGCPQVN	62
QY	61	LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKLQADDNGDPGRNGJLLPSTG	120
DB	63	LPPLTIQLPKQFGRIEEVFKEVQNLKEIVNSLKKTKCQDCKLQADDNDRDPRNGLLSPTG	122
QY	121	APGEVSDNRVRELESEVNKLSSSELKNVAKKEEINVJHGRLEKLNLMNMNIENYVDSKVANL	180
DB	123	APGEADDSRVRELENEVNKLSSDLKNVAKKEEIDGLQGRLEKLSLVNMNIENYVDSKVANL	182

QY 181 TPTVNSLDGKC-SKCPSCQEQIQRPVQHLYKDCSDYYAIGKSSSEYRVTPTDPKSSFE 239
DB 183 TFAVNSLDGKSSRCPSCQEQIQRPVQHLYKDCSEYTTIGKSSSEYRVTPTDPKSSFE 242
QY 240 VYCDMETMGGGWTVLQARLDGSTNFTRTWQDYKAGFGLNRPREFWLGNDKIHLLTKSKEMI 299
DB 243 VYCDMETMGGGWTVLQARLDGSTNFTRTWQDYKAGFGLNRPREFWLGNDKIHLLTKSKEMI 302
QY 300 LRIDLEDNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTTT 359
DB 303 LRIDLEDNGIKLYALYDHFYVANEFLKYRLHIGNYNGTAGDALRFNKHYNHDLKFFTTT 362
QY 360 DKONDRYPSGNCGLYYSSGWWFADCLSANLNGKYHQYRGVNGIFWGTWPGVSEAHPG 419
DB 363 DRONDRYPSGNCGLYYSSGWWFADCLSANLNGKYNQYRGVNGIFWGTWPGVSEAHPG 422
QY 420 GYKSSFKKAKMMIRPKHFKP 439
DB 423 GYRSSFKEAKMMIRPKYFXP 442
RESULT 2
Q9EPT7
ID Q9EPT7 PRELIMINARY; PRT; 357 AA.
AC Q9EPT7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Prothrombinase FGL2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Rychlik D.F., Chien E., Phillippe M.;
RT "FGL2 Expression in the Sprague-Dawley Rat.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF323608; AAG42269.1; -.
DR HSSP; P02671; 1F2D.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN AG_C DOMAIN; 1.
SQ SEQUENCE 357 AA; 40966 MW; 31183D29A02EB5A9 CRC64;
Query Match 69.3%; Score 1647; DB 11; Length 357;
Best Local Similarity 82.7%; Pred. No. 2e-102;
Matches 302; Conservative 24; Mismatches 31; Indels 8; Gaps 1;
QY 75 EEVFKVQNLKEIVNSLKKSCQCKLQADDNGDPGRNGLLPLSTGAFGEVGNRVPELE 134
DB 1 MEEVLKEVRLQEAVDLSLKKSCQCKLQADEHFDPPGNG-----AETAEDNRVQELE 52
QY 135 SEVNLKSLSEKNAKEEINVLRLEKLNLVNANNIENYVDSKVANLTFVWNSLDGKSKC 194
DB 53 SQVNLKSLSEKNAKEEIQGLGRLESILQVNNIENYVDSKVANLTSVNSLDSKCFKC 112
QY 195 PSQEQIQRSPVQHLYKDCSDYYAIGKSSSEYRVTPTDPKSSSEYVCDMETMGGGWTVL 254
DB 113 PSQEHQNPVQHLYKDCSDYYVLGKRSGTYRVTPTDHRNSSSEYVCDMETTGGGWTVL 172
QY 255 QARLDGSTNFTRTWQDYKAGFGLNRPREFWLGNDKIHLLTKSKEMILRIDLEDNGVELYA 314
DB 173 QARLDGSTNFTRTWQDYKAGFGLNRPREFWLGNDKIHLLTKSKEMILRIDLEDNGVELYA 232
QY 315 LYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTTPTDKNDRYPSGNCGLY 374
DB 233 VYDQFYVANEFLKYRLHVGNYNGTAGDALRFNRHYNHDLRFFTTPTDRENDRYPSGNCGLY 292
QY 375 YSSGWWFADCLSANLNGKYHQYRGVNRNGIFWGTWPGVSEAHPGSKSPKAKMMIRP 434

DB 293 YSSGWWFADCLSANLNGKYNQYKGVNRNGIFWGTWPGVSCAHPGGYKFSKAKMMIRP 352
QY 435 KHFKP 439
DB 353 KSFKP 357
RESULT 3
Q8WWE4
ID Q8WWE4 PRELIMINARY; PRT; 148 AA.
AC Q8WWE4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to fibrinogen-like 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017813; AAH17813.1; -.
SQ SEQUENCE 148 AA; 16561 MW; 456B7DBA37C07F21 CRC64;
Query Match 31.5%; Score 750; DB 4; Length 148;
Best Local Similarity 99.3%; Pred. No. 7.1e-43;
Matches 144; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKLANWYWLSSAVLATYGFLLVANNETEIEIKDERAKDVCPVRLESRGKCEAGECPYQVS 60
DB 1 MKLANWYWLSSAVLAAYGFLVANNETEIEIKDERAKDVCPVRLESRGKCEAGECPYQVS 60
QY 61 LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKLQADDNGDPGRNGLLPLSTG 120
DB 61 LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKLQADDNGDPGRNGLLPLSTG 120
QY 121 APGEVGNRVRELESEVNLKLSSELK 145
DB 121 APGEVGNRVRELESEVNLKLSSELK 145
RESULT 4
Q9NDQ1
ID Q9NDQ1 PRELIMINARY; PRT; 652 AA.
AC Q9NDQ1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Fibrinogen-like protein.
GN Cl-FIBRN.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
RA Saitoh N.;
RT "Characterization of Brachyury downstream notochord genes in the Ciona
intestinalis embryo.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036849; BAB00626.1; -.
DR HSSP; P02671; 1F2D.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN AG_C DOMAIN; 1.
SQ SEQUENCE 652 AA; 73252 MW; A492BA325162F0EC CRC64;

Db 192 -----EQKVLMEGHSSEQLSKKEQKDELQVLVSKQSSVIDELEKLVATVNN--SL 243
QY 173 VDSKVANLTFVWNSLDGKCKSCPSQFQIQSRPVCHLIYKDCSDYYAIGKRSSSEYRVTPD 232
Db 244 LQKQCHLMTVNSLLTMMSSPNSKSSVAIRKEEQTTFRDCAEIFKSLTTSGLTSTLTFP 303
QY 233 PKNSSFEVYCDMETYGGGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLL 292
Db 304 NSTEEIKAYCDMEVGGGWTVIQHRDGSVDFTWKEVKEGFGSPLGEYWLGNFVSQSL 363
QY 293 TKSKEMLRIDLEDENGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHD 352
Db 364 TGGHRYVLKIQLKDWEGNEAHSLYDHFYLAGEESNYRHLTGLTGTAGKISS-SQPGSD- 422
QY 353 LKFFTTDPKDNDRYPGNGCLYSSGWMFADCLSANLNGKYVHQYRGVR-NGIFWGTWP 411
Db 423 ---FSTKDSNDKIC-KCSQMLSGWMFADCGPSNLNGQYYPQKQNTNKFNGIKWYYWK 478
QY 412 GVSEAHPGGYKSSFKEAKMMIRPKHF 437
Db 479 G-----SGY--SLKATTMMIRPADF 496

RESULT 7

Q9DER1
ID Q9DER1 PRELIMINARY; PRT: 407 AA.
AC Q9DER1;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE Angiopoietin-2B.
GN ANGIOPOIETIN-2B.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422311; PubMed=10364717;
RA Mezquita J., Mezquita P., Montserrat P., Mezquita B., Francone V.,
RA Vilagrasa X., Mezquita C.;
RT "Genomic structure and alternative splicing of chicken angiopoietin-
2";
RL Biochem. Biophys. Res. Commun. 275:643-651(2000).
DR EMBL; AJ289778; CAC08175.1; -.
DR HSSP; P02671; IFZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 407 AA; 46687 MW; 8963BACAB8C7A41C CRC64;

Query Match 22.0%; Score 524; DB 13; Length 407;
Best Local Similarity 38.3%; Pred. No. 3.4e-27;
Matches 125; Conservative 45; Mismatches 120; Indels 36; Gaps 10;

QY 127 DNRVRELESEVKNLSSELKNAKEEINVLR-----LEK-LNLVNMNI-----ENVV 173
Db 203 EKRVLMEEDKHTLQLKSIKDEKQQLVLRQNSIIIELEKQLVATVNNSVLQKQHDL 162
QY 174 DSKVAN-LTFVNSLDGKCKSCPSQEQIQSRPVQHLIYKDCSDYYAIGKRSSSEYRVTPD 232
Db 163 MERVHNLTMISTPNSAKNFIAKEEQIS-----FKDCBAEAFKSLTTSGLTSTLTFP 214
QY 233 PKNSSFEVYCDMETYGGGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLL 292
Db 215 NSAQEKKAYCDMESNGGGWTVLQRREDGSVDFTWKEVKEGFGDPAGEYWLGNFVSQSL 274
QY 293 TKSKEMLRIDLEDENGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHD 352
Db 275 TNOKRYVLKIILKDWEGNEAYTLTDQFYLANEEQKYRIH-KGLTGTAGK-SSISQPGND- 333

QY 353 LKFFTTDPKDNDRYPGNGCLYSSGWMFADCLSANLNGKYVHQYRGVR-NGIFWGTWP 411
Db 334 ---FSTKDAADKIC-KCSQMLTGGWMFADCGPSNLNGMYYPPLRQNNNKFKGIKWYYWK 389
QY 412 GVSEAHPGGYKSSFKEAKMMIRPKHF 437
Db 390 G-----SGY--SLKATTMMIRPADF 407

RESULT 8

Q9DER0
ID Q9DER0 PRELIMINARY; PRT: 441 AA.
AC Q9DER0;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE Angiopoietin-2C.
GN ANGIOPOIETIN-2C.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422311; PubMed=10964717;
RA Mezquita J., Mezquita P., Montserrat P., Mezquita B., Francone V.,
RA Vilagrasa X., Mezquita C.;
RT "Genomic structure and alternative splicing of chicken angiopoietin-
2";
RL Biochem. Biophys. Res. Commun. 275:643-651(2000).
DR EMBL; AJ289779; CAC08176.1; -.
DR HSSP; P02671; IFZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 441 AA; 50472 MW; DC98127FE0CE34E2 CRC64;

Query Match 22.0%; Score 524; DB 13; Length 441;
Best Local Similarity 38.3%; Pred. No. 3.8e-27;
Matches 125; Conservative 45; Mismatches 120; Indels 36; Gaps 10;

QY 127 DNRVRELESEVKNLSSELKNAKEEINVLR-----LEK-LNLVNMNI-----ENVV 173
Db 137 EKRVLMEEDKHTLQLKSIKDEKQQLVLRQNSIIIELEKQLVATVNNSVLQKQHDL 196
QY 174 DSKVAN-LTFVNSLDGKCKSCPSQEQIQSRPVQHLIYKDCSDYYAIGKRSSSEYRVTPD 232
Db 197 MERVHNLTMISTPNSAKNFIAKEEQIS-----FKDCBAEAFKSLTTSGLTSTLTFP 248
QY 233 PKNSSFEVYCDMETYGGGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLL 292
Db 249 NSAQEKKAYCDMESNGGGWTVLQRREDGSVDFTWKEVKEGFGDPAGEYWLGNFVSQSL 308
QY 293 TKSKEMLRIDLEDENGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHD 352
Db 309 TNOKRYVLKIILKDWEGNEAYTLTDQFYLANEEQKYRIHLSLTGTAGKISS-SQPGND- 367
QY 353 LKFFTTDPKDNDRYPGNGCLYSSGWMFADCLSANLNGKYVHQYRGVR-NGIFWGTWP 411
Db 368 ---FSTKDAADKIC-KCSQMLTGGWMFADCGPSNLNGMYYPPLRQNNNKFKGIKWYYWK 423
QY 412 GVSEAHPGGYKSSFKEAKMMIRPKHF 437
Db 424 G-----SGY--SLKATTMMIRPADF 441

RESULT 9

Q9DER2
ID Q9DER2 PRELIMINARY; PRT: 493 AA.
AC Q9DER2;

Db 17 GREGWALESESCLE---QVR-LRAQVHQLETRVKKQQTMIACLHEK--EVQFLDKGS 69
QY 169 IENYVDSKVANLTFVNVNSLDGKCKSPSQEQIQSRPVQHLIYKDCSDYYAIGKRSSEYR 228
Db 70 ENSFID-----LGGKQ-----YADCSEIYNQGFQKSGEYK 100
QY 229 VTPDPKNSFEVYCDMETGSGWTVLQARLDGSGINFTWQDYKAGFGLN---RREFWL 285
Db 101 IKPLQSLAEFSVYCDMSD-GGGWTVIQRSDGSENFNGWYDENGFGNFVQNGEYWL 159
QY 286 NDKIHLLTKSKEMI,LRIDLEDFNGVELYALYDQFYVANEFLKYRLHVGNVNGTAGDALRF 345
Db 160 NKVINLLT,QQDYTLKIDLTDFEKNSSFAQYQSFYVGDGKKSFYELNIGESGTAGDSLSG 219
QY 346 NKH-----YNHDLKFFTPDKDNDRYPSGNCGLYVSSGKWFADCLSANLNGKYYHQYR 399
Db 220 TFHPEVQWAWASHQRM,KFSTWDRDNDNY-QGNCAEEQSGWFWNRCHSANLNGVYRGYSR 278
QY 400 G-VRNGIFWGTWPGVSEAHPGGYKSSFKKAKMMIRPKFKP 439
Db 279 AETDNGVWVYTW-----HGWWSLKSVVMKIRPSDFIP 311

RESULT 12
Q8NG32 Q8NG32 PRELIMINARY; PRT; 312 AA.
AC Q8NG32;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE LFIRE1.
GN LFIRE1.
CS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Yan J., Yu Y., Wang N., Xu Y.;
RT "LFIRE-1, a liver-specific expressing gene on human chromosome 8p22,
RT is frequently down-regulated and functions as tumor suppressor in
RT human hepatocellular carcinoma";
RI Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CR EMBL; AF168954; AAU88911.1; -
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 312 AA; 36377 MW; 7330F3D55AC5D619 CRC64;

Query Match 21.3%; Score 506; DB 4; Length 312;
Best Local Similarity 44.6%; Pred. No. 3.9e-26;
Matches 107; Conservative 25; Mismatches 88; Indels 20; Gaps 6;
QY 210 YKDCSDYYAIGKRSSEYRVTPDPKNVSSFEVYCDMETMGCGWTVLQARLDGSGTNFTRWQ 269
Db 80 YADCSEIFNDGYKLSGFYKIKPLQSPAEFSVYCDMSD-GGGWTVIQRSDGSENFNGWK 138
QY 270 DYKAGFGNL---RREFWLGNDKIHLLTKSKEMILRIDLEDFNGVELYALYDQFYVANEFL 326
Db 139 DYENGFGNFVQKHGEYWLGNLKNLHFLTTQEDYTLKIDLAQFEKNSRYAQYKQKVGDEKN 198
QY 327 KYRLHVGNVNGTAGDALRFNKH-----YNHDLKFFTPDKDNDRYPSGNCGLYSSGW 380
Db 199 FYELNIGESGTAGDSLAGNFHPEVQWAWASHQRMKFTWDRDHDNY-EGNCABEDQSGW 257
QY 381 FDACLSANLNGKYYHQYRG-VRNGIFWGTWPGVSEAHPGGYKSSFKKAKMMIRPKFKP 439
Db 250 FNRCHSANLNGVYSGPYTAKTNGIVWYTW-----HGWWSLKSVVMKIRPNDIFIP 309

RESULT 13
Q95841 Q95841 PRELIMINARY; PRT; 491 AA.
AC Q95841;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Angiopoietin Y1 (DJ595C2.2) (Angiopoietin-related protein 1 precursor).
GN DJ595C2.2 OR ARP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=99148929; PubMed=10025962;
RA Kim I., Kwak H.J., Ahn J.E., So J.N., Liu M., Koh K.N., Koh G.Y.;
RT "Molecular cloning and characterization of a novel angiopoietin family protein, angiopoietin-3";
RT FEBS Lett. 443:353-356(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Cobley V.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Ota T., Nishikawa T., Kawai Y., Suzuki Y., Ishii S., Saito K., Yamamoto J., Sugano S., Isogai T.;
RT "HRI human cDNA sequencing project";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Ito Y., Oike Y., Yasunaga K., Matsumoto S., Ota T., Nishikawa T., Kawai Y., Isogai T., Hamada K., Saito Y., Miyata K., Masuho Y., Suda T.;
RT "Molecular cloning and characterization of novel angiopoietin-related protein (ARP4)";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CR EMBL; AF107253; AAD19608.1; -
CR EMBL; AL355520; CAC13169.1; -
CR EMBL; AB056476; BAB40691.1; -
CR HSSP; P02671; 1PZD.
DR Genew; HGNC:489; ANGPTL1.
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 491 POTENTIAL.
SQ SEQUENCE 491 AA; 56719 MW; 3C4DB8DEF6CF7E99 CRC64;

Query Match 21.2%; Score 505; DB 4; Length 491;
Best Local Similarity 26.9%; Pred. No. 8.1e-26;
Matches 131; Conservative 81; Mismatches 137; Indels 138; Gaps 15;
QY 47 GXCEEAGECPYQVSLPPLTIQIP-----KQFSRIEEVF--KEVQNLKEIVNSLKKSCQ 97
Db 42 GK-EBAKKCAYTFLLVPEQRITGPICVNTKGQDASTIKDMITRMDLENLKDVLSPQKREID 100
QY 98 DCKLOADDNGDPGRNGSLLPSTGAPGEVDNRVRELSSEVNKLSSSELKNAKEEINVLHGR 157
Db 101 VLQLVVDVJGN-----IVNEVKLLRKESRNMNSRVQLYMQ 136
QY 158 L-----EKLNLVNMNNIENYV-----DSKANLTFVNS-----L 187
Db 137 LLHEIIRKRDNSLELSQLENKILNVTTTEMLKMATRYRELEVKYASLTDLVNNSQSVMTLL 196
QY 188 DQKCKSPSQEQICSRP-----VQHL----- 208

Db 197 EQCLRIFSRQDTHVSPBLVQVVPQHIPNSQQYTPGJLGGNEIQRPDPGYPRDLMPPDLA 256
QY 209 -----IYKQSDYYAIGKRSETYAVTPDPKNSSEFYCDMETMGG 249
Db 257 TSPTKSPFKIPPVTFINEGPFKDCQCAKEAGHSVSGIYMIKPENSGPMQLWCENSLOPG 316
QY 250 GWTVLQARLDGSTNFTRWQDYKAGFGLNRRREFWJGNDKIHLITKSKEMILRIDLEDFNG 309
Db 317 GWTVIQKRTDGSVNFRRWENYKKGFGNIDGEYWLGLNIYMLSNQCNKYKLLIELEDWSD 376
QY 310 VELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRFNKHYNHDLKFETTPDKNDORYPSG 369
Db 377 KKVYAEYSSFRLEPESEFYRLGLGTQYQGNAGDSMMW-----HNGKQFTTLDRKDWY-AG 430
QY 370 NCGLYSSGWWFDCLSANLNGKYVH-CXYRGV-RNGIFWGTWPGVSEAHFGGYKSSFKE 427
Db 431 NCAHFHKGWYWNACAHNSNLNGVWYRGGHYRSKHQDGFPAEYRG-----GSY--SLRA 492
QY 428 AQMMIRP 434
Db 483 VQMMIKP 489
RESULT 14
Q9HBP3 PRELIMINARY; PRT; 244 AA.
AC Q9HBP3;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
DE Hypothetical protein.
CS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.X., Qian J.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
growth.";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBSJ databases.
RL EMBL; AF218015; AAG17257.1; -.
DR HSSP; PC2671; 1F2D.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 244 AA; 27825 MW; E9D0ACBF87A71139E CRC64;
Query Match 20.9%; Score 496; DB 4; Length 244;
Best Local Similarity 38.9%; Pred. No. 1.3e-25;
Matches 102; Conservative 43; Mismatches 35; Indels 22; Gaps 6;
QY 177 VANLTFVNSLDGKCKSPQEQIQSRPVQHLLIYKDCSDYYAIGKRSETYAVTPDPKNS 236
Db 4 VNLLTNMSTNSKOPTVAKEEQIS-----FRDCAEVFKSGHTTNG-YTLTFPNSTE 55
QY 237 SFEVYCDMETMGGWTVLQARLDGSTNFTRTWQDYKAGFGLNRRREFWJGNDKIHLITKSK 296
Db 56 EIKAYCDMEAGGGWTTIQRREDGSDVFQRTWKEYKYVFGNPSGEYWLGNFVSLTNQQ 115
QY 297 EMILRIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRFNKHYNHDLKF 356
Db 116 RYVLKIRLKDWEAGNEAYSLYEHFYLSSSEELNYRIHLKGLTGAGKISSSQPND---F 171
QY 357 TTPDKNDORYPSGNCGLYSSGWWFDCLSANLNGKYVH-QYRGV-RNGIFWGTWPGVSE 415
Db 172 STKGDNDKIC-KCSQMLTGGWFDACGSPSNLNGMYYPQRTNKNKFGIKWYWKG--- 227

QY 416 AHPGGYKSSFKKAKMMIRPKHF 437
Db 228 ---SGY--SLKATTMMIRPADF 244
RESULT 15
Q90Z19 PRELIMINARY; PRT; 513 AA.
AC Q90Z19;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Angiopoietin-1.
GN ANGPT1 OR ANG1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=21391693; PubMed=11500995;
RA Pham V.N., Roman B.L., Weinstein B.M.;
RT "Isolation and expression analysis of three zebrafish angiopoietin
genes.";
RL Dev. Dyn. 221:470-474 (2001).
DR EMBL; AF379602; AAK83347.1; -.
DR ZFIN; ZDB-GENE-010817-1; angpt1.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 513 AA; 58360 MW; 0627777A39847D9B CRC64;
Query Match 20.7%; Score 491.5; DB 13; Length 513;
Best Local Similarity 26.6%; Pred. No. 6.8e-25;
Matches 142; Conservative 86; Mismatches 179; Indels 127; Gaps 16;
QY 8 WJSSAVLATYGLVAVNANETEEIKDERAKDVCVRLESRGKCE---EAGECPYQVSLPPL 64
Db 2 WKGCLFLA--ALLVADCGGVQKTDGLWSSTPKSNSSSGRRYHRIHQGCSTFILPES 59
QY 65 TIQLPKQF-----SRIEEVFKB-VQNKEIVNSLKSKC 96
Db 60 DGNTCREFKSGTAYNANALQRAPQPEADLSNQIKQLEHVMENYTWLQKIENYIKDNM 119
QY 97 QCKLQADDNGDPRNGLLLPSTGAPGEVGDN-----RVRELESEV----- 137
Db 120 KTEMVQLQQSAVHNHTAAML-----EMGTSLLSQTAEQTRKLTDTVETQVLNQTSLREI 172
QY 138 -----NKLSELKNABEINVHLH---GRLE-----KLNLVNMNNIENVDS 175
Db 173 QLENSLSTNKLKQLMIQINEINKIHDKNGFLEEKMQLEDRHROELESRLTEKSDLQA 232
QY 176 KVANLTFVNSLDGKCKSPQEQIQSRPVQHLLI----- 209
Db 233 LVSRQSSVIRELENQLSRATGNSTALQROQQDLMESMRSLLSLCAKDAATAVEPNSTKQA 292
QY 210 -----YKDCSDYYAIGKRSETYAVTPDPKNSSEFYCDMETMGGWTVLQARLDGSTNF 264
Db 293 DEERKFRDCADLYQAGFQKNGVYVYINISQETK-KVYCVMESAGGGWTVIQKREDGVDF 351
QY 265 TRTWQDYKAGFGLNRRREFWJGNDKIHLITKSKEMILRIDLEDFNGVELYALYDQFYVANE 324
Db 352 QXTWKYKYGFGSVSGEHWLGNFVHVLTNQRQHGLRVELSDWDGHOAFSQYDSFHDSE 411
QY 325 FLKYRLHVGNYNGTAGDALRFNKHYNHDLKFETTPDKNDORYPSGNCGLYSSGWWFDCAC 384
Db 412 KQYRLFLKTHSGTAGRQSSLAHV-GAD---FSTKDVNDNC-TCKCALMLSGGWWYDAC 466
QY 385 LSAHLNGKYVHQ-KYRGVNRNGIFWGTWPGVSEAHFGGYKSSFKKAKMMIRPKHF 437
Db 467 GPSNLNGVYRQGGHVGGKFNIGIKWHYFKGPS-----YSLSRSTVMIRGAJF 512

Search completed: November 5, 2003, 16:47:34
Job time : 61.0176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 15:56:26 ; Search time 86.0595 Seconds
(without alignments)
809.684 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MKLANWYWLSSAVLATYGPL.....GYKSSFEAKMMIRPKHKP 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2378	100.0	439	AAW88235	Human prothrombina
2	1853.5	77.9	432	AAW88236	Mouse prothrombina
3	524	22.0	314	AAR65759	Rat hepatic parench
4	522	22.0	496	AAW01411	Human TIE-2 ligand
5	522	22.0	496	AAR94635	Human TIE-2 ligand
6	522	22.0	496	AAW47532	Human TIE-2 ligand
7	522	22.0	496	21	Human angiotensin
8	522	22.0	496	21	Human angiotensin
9	522	22.0	496	21	Human angiotensin

10	522	22.0	496	23	AAU77344	Amino acid sequenc
11	522	22.0	496	24	ABU07855	Human angiotensin
12	522	22.0	496	24	ASP58064	Human angiotensin
13	522	22.0	572	22	ABG06771	Novel human diagno
14	522	22.0	1033	22	ABG11614	Novel human diagno
15	521.5	21.9	498	19	AAW47528	Amino acid sequenc
16	520	21.9	444	23	AAU77252	Human angiotensin
17	517.5	21.8	499	19	AAW47529	Amino acid sequenc
18	506.5	21.3	286	23	ABG80360	Human TIE-2 ligand
19	506.5	21.3	491	21	AAW48000	Protein encoded by
20	506	21.3	312	17	AAR94317	Hepatocyte prolif
21	506	21.3	339	21	AAW43637	Human cancer assoc
22	505.5	21.3	312	17	AAR94316	Hepatocyte prolif
23	505	21.2	491	20	AAW26196	Human zapo3 protei
24	505	21.2	491	20	AAW05397	Human TIE ligand N
25	505	21.2	491	21	AAW47396	Angiotensin-like
26	505	21.2	491	21	AAW24393	Human PRO188 prote
27	505	21.2	491	21	AAW69483	Amino acid sequenc
28	505	21.2	491	21	AAW68761	An angiotensin-re
29	505	21.2	491	22	AAU12310	Human PRO188 polyp
30	505	21.2	491	22	AAW60340	Human neovasculari
31	505	21.2	491	22	AAW53067	Human angiogenesis
32	505	21.2	491	23	ABG80356	Human TIE ligand N
33	505	21.2	491	23	AAE19826	Human TIE ligand N
34	505	21.2	491	24	ABU66709	Human PRO polypept
35	505	21.2	491	24	ABU66984	Human secreted/tra
36	505	21.2	491	24	ABU59789	Novel secreted and
37	505	21.2	491	24	ABG72425	Human angiotensin
38	504.5	21.2	312	15	AAR65760	Human hepatic pare
39	500	21.0	491	21	AAW44841	Human secreted pro
40	495	20.8	282	24	ABR48466	Human Frangiopogen
41	490	20.6	689	21	AAW96780	Ang-2-FD-FC-FD fus
42	489	20.6	346	20	AAW23736	Human AR-1 protein
43	489	20.6	346	20	AAW05399	Human TIE ligand N
44	489	20.6	346	21	AAW90400	Human angiotensin
45	489	20.6	346	21	AAW33432	Human PRO356 prote

ALIGNMENTS

RESULT 1
AAW89235
ID AAW89235 standard; Protein; 439 AA.

AC AAW89235;
XX
CT 15-MAR-1999 (first entry)
XX Human prothrombinase Fgl2 protein.
DE Prothrombinase; hfgl2; Fgl2; human; immune coagulation; antibody;
KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;
KW gastrointestinal disease; foetal loss; therapy; vaccine.
XX Homo sapiens.

Key	Location/Qualifiers
Modified-site	179..183
Modified-site	/label= Asn is N-glycosylated
Modified-site	235..238
Modified-site	/note= "Asn is N-glycosylated"
Modified-site	262..265
Modified-site	/note= "Asn is N-glycosylated"
Modified-site	336..337
Modified-site	/note= "Asn is N-glycosylated"
Domain	213..439
Peptide	/note= "fibrinogen related domain"
Peptide	364..378
Peptide	/note= "epitope (Claim 4)"
WC9851335-A1.	

```
PD 19-NOV-1998.
XX
PF 15-MAY-1998; 98WO-CA00475.
XX
FR 10-OCT-1997; 97US-0061684.
PR 15-MAY-1997; 97US-0046537.
XX
PA (LEVY/) LEVY G.
XX
FI Levy G;
XX
DR WPI; 1999-059687/05.
DR N-PSDB; AAV84139.
XX
PF Modulating immune coagulation - by using Fgl2 antibodies and
PF compounds, used to treat conditions including graft rejection and
PF foetal loss
XX
PS Claim 8; Page 66-67; 105pp; English.
XX
PD This is the amino acid sequence of human prothrombinase Fgl2, as
CC predicted from hfgl2 DNA (see AAV84139). Fgl2 is a 70 kDa
CC transmembrane serine protease that has immune procoagulant activity.
CC The invention provides a method for inhibiting immune coagulation by
CC inhibiting the activity or expression of Fgl2. The method can be
CC used in vivo to treat a condition which requires a reduction in
CC immune coagulation such as bacterial and viral infections, cancer,
CC glomerulonephritis, a number of gastrointestinal diseases,
CC allograft and xenograft rejection and foetal loss. An Fgl2-specific
CC antibody, an Fgl2 antisense oligonucleotide, or a substance that
CC affects prothrombinase activity of a Fgl2 protein may be used to
CC treat a condition requiring a reduction in procoagulant activity.
CC A vaccine containing an Fgl2 protein or peptide is used for
CC prevention of graft rejection or foetal loss (claimed).
XX
SQ Sequence 439 AA;
Query Match 100.0%; Score 2378; DB 20; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.1e-205;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLANWYWLSSAVLATYGFVLVANNETEIEIKDERAKVCFVRLESRGKCEAGECPYQVS 60
Db 1 MKLANWYWLSSAVLATYGFVLVANNETEIEIKDERAKVCFVRLESRGKCEAGECPYQVS 60
QY 61 LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKQCADNDGDPGRNGLLLPSTG 120
Db 61 LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKQCADNDGDPGRNGLLLPSTG 120
QY 121 APGEVCGNVRVRELESEVNKLSSSELKNAKEEINVLHGRLEKLNLYANNNIENYVDSKVANL 180
Db 121 APGEVCGNVRVRELESEVNKLSSSELKNAKEEINVLHGRLEKLNLYANNNIENYVDSKVANL 180
QY 181 TFVNSLDGKCKSCPSQEQICSRPVQHLIYKDCSDYYAIGKRSSETYRVTPDPKNSFEV 240
Db 181 TFVNSLDGKCKSCPSQEQICSRPVQHLIYKDCSDYYAIGKRSSETYRVTPDPKNSFEV 240
QY 241 YCDMETMGSGWTVLQARLDGSTNFTETWQDYKAGFGNRRERWLGNDKIHLLTKSKEMIL 300
Db 241 YCDMETMGSGWTVLQARLDGSTNFTETWQDYKAGFGNRRERWLGNDKIHLLTKSKEMIL 300
QY 301 RIDLEDENGVELYALYDQFYVANEFLLKYLRLHVGNNGTAGDALRFNKHYNHDLKFFTFPD 360
Db 301 RIDLEDENGVELYALYDQFYVANEFLLKYLRLHVGNNGTAGDALRFNKHYNHDLKFFTFPD 360
QY 361 KQNDRYPSGNCGLYSSGWWFDAQLSANLNGKYYHQYRGVNRNIGIFWGTWPGVSEAHFPGG 420
Db 361 KQNDRYPSGNCGLYSSGWWFDAQLSANLNGKYYHQYRGVNRNIGIFWGTWPGVSEAHFPGG 420
QY 421 YKSSFKEAKMMIRPKHFKP 439
Db 421 YKSSFKEAKMMIRPKHFKP 439
```

```
RESULT 2
AAW88236
ID AAW88236 standard; Protein; 432 AA.
XX
AC AAW88236;
XX
DT 15-MAR-1999 (first entry)
XX
DE Mouse prothrombinase Fgl2 protein.
XX
KW Prothrombinase; Fgl2; mouse; immune coagulation; antibody;
KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;
KW gastrointestinal disease; foetal loss; therapy; vaccine.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Modified-site 172..174
FT /label= Asn is N-glycosylated
FT Modified-site 228..231
FT /note= "Asn is N-glycosylated"
FT Modified-site 256..259
FT /note= "Asn is N-glycosylated"
FT Modified-site 323..325
FT /note= "Asn is N-glycosylated"
FT Domain 213..439
FT /note= "fibrinogen related domain"
XX
WQ9851335-A1.
PN 19-NOV-1998.
PD 15-MAY-1998; 98WO-CA00475.
XX
PR 10-OCT-1997; 97US-0061684.
PR 15-MAY-1997; 97US-0046537.
XX
PA (LEVY/) LEVY G.
XX
PI Levy G;
XX
DR WPI; 1999-059687/05.
DR N-PSDB; AAV84140.
XX
PT Modulating immune coagulation - by using Fgl2 antibodies and
PT compounds, used to treat conditions including graft rejection and
PT foetal loss
XX
PS Claim 8; Page 70-71; 105pp; English.
XX
CC This is the amino acid sequence of mouse prothrombinase Fgl2, as
CC predicted from fgl2 DNA (see AAV84140). Fgl2 is a 70 kDa
CC transmembrane serine protease that has immune procoagulant activity.
CC The human Fgl2 amino acid sequence is given in AAW88236. The
CC invention provides a method for inhibiting immune coagulation by
CC inhibiting the activity or expression of Fgl2. The method can be
CC used in vivo to treat a condition which requires a reduction in
CC immune coagulation such as bacterial and viral infections, cancer,
CC glomerulonephritis, a number of gastrointestinal diseases,
CC allograft and xenograft rejection and foetal loss. An Fgl2-specific
CC antibody, an Fgl2 antisense oligonucleotide, or a substance that
CC affects prothrombinase activity of a Fgl2 protein may be used to
CC treat a condition requiring a reduction in procoagulant activity.
CC A vaccine containing an Fgl2 protein or peptide is used for
CC prevention of graft rejection or foetal loss (claimed).
XX
SQ Sequence 432 AA;
```

```
Query Match 77.9%; Score 1853.5; DB 20; Length 432;
Best Local Similarity 77.7%; Pred. No. 3.9e-158;
Matches 341; Conservative 42; Mismatches 49; Indels 7; Gaps 4;
```

QY	1	MKLANWYWLSSAVLATYGFVLVANNETEETIKDERAKDVCPVRLESRGKEEAGECPYQVS	60
DB	1	MRLPGWLWSSAVLAACR-AVEEHNLTGLELDASQAACPARLEGSGRCEGSQCPFLT	58
QY	61	LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKLQADCGDPGRNGLLLPSTG	120
DB	59	LPPLTIQLPRQLGSMEEVLEKVRTLEAVIDSLKKSCQCKLQADCHRDGNG---	114
QY	121	APGEVGDNRVRELESEVNLSSSELXNAKEEINVLEHGRLEKLVNMNN-ENYVDSKVANL	180
DB	115	AS-TAEDSRVQELSCVNLKLSSELXNAKDQIQGLGRLETLHVNMMNNENYVDNKVANL	173
QY	181	TFVVNSLDGKCKSPSQEQIQSRPVQHLIYKQCSDYAIGKRSSSEYRVTPDPKNSSFEV	240
DB	174	TVVNSLDGKCKSPSQEHMQSPVQHLIYKQCSDHVJLGRSSGAYRVTPDHRNSSFEV	233
QY	241	YCDMETMGGGWTVLQARLDGSTNFTRTWQDYKAGFGLRREFWLGNDKZHLHTKSKEHIL	300
DB	234	YCDMETMGGGWTVLQARLDGSTNFTREWKDYKAGFGLNLEREFWLGNDKZHLHTKSKEHIL	293
QY	301	RIDLEDENGVELYALYDQFYVANEFLKYRLHVGNVNYNGTAGDALRPNKHYNHDLKFPTPD	360
DB	294	RIDLEDENGTLTYALYDQFYVANEFLKYRLHIGNVNGTAGDALRFRSHYNHDLRFFTTPD	353
QY	361	KNDRIYPSGNCGLYYSSGWWFDACLSANLNGKYYHQYRGVRNGIFWGTWPGVSEAHPGG	420
DB	354	RNDRIYPSGNCGLYYSSGWWFDSCLSANLNGKYYHQYKGVRENGIFWGTWPGINCAQPPGG	413
QY	421	YKSSFKKAKMMIRPKHFKP	439
DB	414	YKSSFKQAQKMMIRPKHFKP	432

RESULT 3
AAR65759
ID AAR65759 standard; protein: 314 AA.

DT	25-MAR-2003 (updated)
DT	17-MAY-1995 (first entry)

DE Rat hepatic parenchymal cell growth factor.

KW Hepatic parenchymal cell growth factor; HPGF; liver diseases;
 KW liver cancer; cirrhosis.

C5 Rattus rattus.

FT	Key	Location/Qualifiers
FT	Peptide	1..24
FT		/label= sig peptide

PN WO9421678-A1.

xx 29-SEP-1994.

22-MAR-1994: 94WO-JP00455.

23-MAR-1993: 93JP-0063905.

(TAIS) TAISHO PHARM CO LTD.

PI Hanada K, Hara H, Matsuki Y, Shindo S, Yoshimura H;

DR WPI; 1994-316940/39.

DR N-PSDB; AAC77818.

PR: Hepatic parenchymal cell growth promoter peptide - is isolated from human or animal liver cell or produced by recombinant techniques and used for therapy of liver diseases

PS Claim 3; Page 34; 47pp; Japanese.

[illegible]

RESULT 4

AAW01411

ID AAW01411 standard; Protein; 496 AA.

AC AAW01411:

11-FEB-1997 (first entry)

Human TIE-2 ligand 2.

TIE-2 ligand 2: tyros

KW leucopaemia; thrombocytopaenia; anaemia; angiodenesis; tumour;
KW leukopenia; thrombocytopenia; anemia; angiogenesis; tumor;
KW receptor antagonist; neovascularisation; wound healing; ischaemia
KW arteriosclerosis; inflammation; diagnosis; therapy.
KW

xx Homo sapiens.

XX
PN
W09637598-A1

10-OCT-1995

AA 96WO-US049C6

06-CCT-1995: 95WC-US12935

XX
XX
00-4FK-1333;
3303-0418333;

OF (REF-) REGENERON PHARM. INC.

Aldrich TH, Bruno J, Davis S, Goldfarb M, Jones PF;

PI Maisonnier PC, Radziejewski C, Yancopoulos GD; Maisnier PC, Radziejewski C, Yancopoulos GD;

DR. WFI; 1996-465021/46.

DR N-PSDB; AAT44321.

TIE-2 agonists and antagonists and related DNA - useful for


```
Db 375 LKDWEGNEAYSLEYEHFYLSSSEELNRYIHLKGLTGAGKISSISQPGND-----FSTKDGDN 430
QY 364 DRYPSGNCGLYYSSGWWFDACLSANLNGKYYHCKYRGVR-NGIFWGTWPGVSEAHFPGGYK 422
Db 431 DKCIC-KCSQMLTGGWWFDACGPNLNGMYYPQONTNKFNGIKWYYWKG-----SGY- 482
QY 423 SSFKEAKMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 6
AAW47532
ID AAW47532 standard; Protein; 496 AA.
XX
AC AAW47532;
XX
DT 09-SEP-1998 (first entry)
XX
DE Human TIE-2 ligand 2 from clone pBluescript KS.
XX
KW Chimeric TIE ligand 2N1CiF; TIE-2 ligand; neovascularisation;
tumour; human.
XX
OS Homo sapiens.
XX
PN WO9805779-A1.
XX
FD 12-FEB-1998.
XX
PF 01-AUG-1997; 97WO-US13557.
XX
PR 25-OCT-1996; 96JS-0740223.
PR 02-AUG-1996; 96US-0022999.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Davis S, Yancopoulos GD;
XX
DR WPI; 1998-145615/13.
DR N-PSEB; AAV18619.
XX
PT Modified human Tie-2 receptor ligand(s) - useful for promoting wound
healing
PT
XX
PS Example 8; Fig 6; 202pp; English.
XX
CC This is the amino acid sequence of the human TIE-2 ligand 2,
used in the method of the invention, involving the production
of TIE-2 ligands which promote healing. The nucleic acids, vectors
and host cells used in the method of the invention are useful for
the recombinant production of the ligands. The ligands, etc. are
useful for blocking blood vessel growth, promoting neovascularisation,
promoting the growth or differentiation of a cell expressing the TIE
receptor, blocking the growth or differentiation of a cell expressing
the TIE receptor and for attenuating or preventing tumour growth in
a human.
XX
SQ Sequence 496 AA;

Query Match 22.0%; Score 522; DB 19; Length 496;
Best Local Similarity 27.9%; Pred. No. 5.2e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVQN 84
Db 26 MDSIGKKQYQVHGSCSYTFLLPENYGRSSSSPYVSNVAVQRPAPLEYDDSVQRLQVLEN 85
QY 85 LKE-----IVNSLKKSCQDCKLOADDNGDPGANGLLPSTGAPGEVGDN---- 128
Db 86 IMENNTQWLKLENYIQDNKKEMVEIQNAVQN-----QTAVNIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSSSELKNAKEEINVLHGR---LEKL 161
```

```
Db 135 TAEQTRKLTQVEAQLNQTTRELEQLLEHSLSTNKKLEQILQOTSEINKLQDKNSFLEKK 194
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
Db 195 VLAMEDKHIIQLQSIKEEKDQLVLVSKNSIIEELEKKIIVTATVNNNSVLQKQCHDLNET 254
QY 184 VNSLDGKSKCPSQEQIQSRPVQHLYKDCSDYYAIGKRSSETYRVTDPKNSSFVEYCD 243
Db 255 VNNLLTMSTNSAKDPTVAXEEQISFRDCAEVFKSGHTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFGLNRREFWLGNDKIHLLTKSKEMILRID 303
Db 315 MEAGGGGWIIQRREDGVDFFQRTWKEYKVGFGNPSGEYWLGNFVSQLTNQORYVLKTH 374
QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKGHYNHDLKFEFTTPDKDN 363
Db 375 LKDWEGNEAYSLEYEHFYLSSEELNRYIHLKGLTGAGKISSISQPGND-----FSTKDGDN 430
QY 364 DRYPSGNCGLYYSSGWWFDACLSANLNGKYYHCKYRGVR-NGIFWGTWPGVSEAHFPGGYK 422
Db 431 DKCIC-KCSQMLTGGWWFDACGPNLNGMYYPQONTNKFNGIKWYYWKG-----SGY- 482
QY 423 SSFKEAKMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 7
AAB28392
ID AAB28392 standard; Protein; 496 AA.
XX
AC AAB28392;
XX
DT 19-FEB-2001 (first entry)
XX
DE Human angiopoietin-2.
XX
KW Human; angiopoietin-2; cytostatic; antiproliferative;
vascular endothelial growth factor; VEGF; antibody; VEGF2 receptor;
blood vessel regression; cancer; vascularised solid tumour.
XX
OS Homo sapiens.
XX
PN WO200064946-A2.
XX
PD 02-NOV-2000.
XX
PF 28-APR-2000; 2000WO-US11367.
XX
PR 28-APR-1999; 99US-0131432.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Thorpe PE, Brekken RA;
XX
DR WPI; 2000-687317/57.
DR N-PSEB; AAC67774.
XX
PT Immunogenic composition for the treatment and diagnosis of cancer
comprises an anti-VEGF (vascular endothelial growth factor) antibody
binding the same epitope as the monoclonal antibody ATCC PTA 1595 -
XX
PS Disclosure; Page 283-285; 298pp; English.
XX
CC The present invention relates to anti-Vascular Endothelial Growth Factor
(VEGF) antibodies that bind to the same epitope as the monoclonal
antibody ATCC PTA 1595 and which significantly inhibit VEGF binding to
the VEGF receptor VEGFR2, without inhibiting VEGF binding to the VEGF
receptor VEGFR1. The present sequence is human angiopoietin-2.
CC Angiopoietin-2 may be operatively attached to the anti-VEGF antibodies of
the present invention. Angiopoietin-2 acts to disturb capillary structure
and is thought to lead to vessel regression. The anti-VEGF antibodies of
```

CC the present invention are useful for the treatment and diagnosis of
CC cancer, especially vascularised solid tumours.

XX Sequence 496 AA;

SQ Query Match 22.0%; Score 522; DB 21; Length 496;
Best Local Similarity 27.9%; Pred. No. 5.2e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEVEFKEVQN 84
DB 26 MDSIGKKQYQVQHGSCTYFLLPEYDNCRSSSPYVSNVAVQRDAPLEYDQSVQRLQVLEN 85
QY 85 LKE-----IVNSLKKSCQDCKLQADDNGDPGRNGLLLPSTGAPGEVGN---- 128
DB 86 IMENNTQWLKLENYIQDNKKEMVEIQNAVN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSELKNAAKEEINVHLGR---LEKL 161
DB 135 TAEQTRKLTQVEAQVNLQNTTRLEQLLEHSLSTNKLEKQILDQTSSEINKLQDNKSFLEKK 194
QY 162 NLV-----NMNNTENYVDSKVANLTF-----V 183
DB 195 VLAMEDKHIIQLQSIKEEKDQLQVLVSKQNSIIELEKKIVTATVNVNSVLQKQHDLMET 254
QY 184 VNSLDGKCKSCPSQEQISRPVQHLYIKDCSDYYAIGKRSSETYRVTDPKNSSEFEVYCD 243
DB 255 VNNLLTMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMILRID 303
DB 315 MEAGGGGTIIQRREDGSDVDFQRTWKEYKVGFGNPSGEYWLGNFVSQLTNQORYVLKIH 374
QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPDKN 363
DB 375 LKDWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAGKISSISQPGND----FSTKGDGN 430
QY 364 DRYPSGNCGLYSSQWFWFDACLSANLNGKYHVKYRGVR-NGIFWGTWPGVSEAHPGGYK 422
DB 431 DKCIC-KCSQMLTGGWFWFDACGSPSNJNGMYFPQRQNTNKFNGIKWYWKG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
DB 483 -SLKATTMMIRPADF 496

RESULT 8
AAV78903 standard; protein; 496 AA.

AAV78903;
19-MAY-2000 (first entry)

Human angiotensin-2 amino acid sequence.

Angiotensin-2; Ang-2; vascular endothelial growth factor; tumour;
vascularisation; angiogenesis; blood vessel maturation; malignant;
benign; binding ligand; cancer; aminophospholipid.

Home sapiens.

WO200002587-A1.

23-JAN-2000.

12-JUL-1999; 99WO-US:5668.

13-JUL-1998; 98JS-0092589.

02-DEC-1998; 98JS-0110600.

(TEXA / UNIV TEXAS SYSTEM.

PI Thorpe BE, Ran S;

XX WPI; 2000-182177/16.

DR N-PSDB; AA792213.

XX Binding ligand for aminophospholipid used in the treatment of
PT vascularised tumours, comprises targeting component and therapeutic
PT agent -

RS Disclosure; Page 257-259; 267pp; English.

XX This sequence represents the human angiotensin-2 (Ang-2) amino acid
CC sequence. Ang-2 is a naturally occurring angiogenesis antagonist when
CC vascular endothelial growth factor (VEGF) levels are low, and generally
CC counteracts the blood vessel maturation and stability mediated by Ang-1.
CC Ang-1 is a naturally occurring angiogenesis agonist, and is a maturation
CC or stabilisation factor, converting immature vessels to mature vessels.
CC Both Ang-1 and Ang-2 are useful in a therapeutic approach to the
CC treatment of vascularised tumours. The invention relates to a binding
CC ligand comprising a targeting agent that binds to an aminophospholipid,
CC linked to a therapeutic agent (e.g. Ang-2). Aminophospholipids are stable
CC and specific markers accessible on the luminal surface of tumour blood
CC vessels. The binding ligand induces coagulation (thrombosis) in tumour
CC vasculature or causes tumour necrosis (possibly by cell- or
CC complement-mediated cytotoxicity and/or apoptosis). The binding ligands
CC are used to treat vascularised tumours, malignant or benign, in animals,
CC especially large tumours.

SQ Sequence 496 AA;

Query Match 22.0%; Score 522; DB 21; Length 496;
Best Local Similarity 27.9%; Pred. No. 5.2e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEVEFKEVQN 84
DB 26 MDSIGKKQYQVQHGSCTYFLLPEYDNCRSSSPYVSNVAVQRDAPLEYDQSVQRLQVLEN 85
QY 85 LKE-----IVNSLKKSCQDCKLQADDNGDPGRNGLLLPSTGAPGEVGN---- 128
DB 86 IMENNTQWLKLENYIQDNKKEMVEIQNAVN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSELKNAAKEEINVHLGR---LEKL 161
DB 135 TAEQTRKLTQVEAQVNLQNTTRLEQLLEHSLSTNKLEKQILDQTSSEINKLQDNKSFLEKK 194
QY 162 NLV-----NMNNTENYVDSKVANLTF-----V 183
DB 195 VLAMEDKHIIQLQSIKEEKDQLQVLVSKQNSIIELEKKIVTATVNVNSVLQKQHDLMET 254
QY 184 VNSLDGKCKSCPSQEQISRPVQHLYIKDCSDYYAIGKRSSETYRVTDPKNSSEFEVYCD 243
DB 255 VNNLLTMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMILRID 303
DB 315 MEAGGGGTIIQRREDGSDVDFQRTWKEYKVGFGNPSGEYWLGNFVSQLTNQORYVLKIH 374
QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPDKN 363
DB 375 LKDWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAGKISSISQPGND----FSTKGDGN 430
QY 364 DRYPSGNCGLYSSQWFWFDACLSANLNGKYHVKYRGVR-NGIFWGTWPGVSEAHPGGYK 422
DB 431 DKCIC-KCSQMLTGGWFWFDACGSPSNJNGMYFPQRQNTNKFNGIKWYWKG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
DB 483 -SLKATTMMIRPADF 496

RESULT 9
AAV78906

ID AAY78906 standard; Protein; 496 AA.
XX
AC AAY78906;
XX
DT 19-MAY-2000 (first entry)
XX
DE Human angiopoietin-2 (Ang-2) amino acid sequence.
XX
KW Human; angiopoietin-2; Ang-2; aminophospholipid; vascularised tumour;
KW unconjugated anti-aminophospholipid antibody; tumour blood vessel marker;
KW cancer; treatment; angiogenesis.
XX
OS Homo sapiens.
XX
PN WO200002584-A2.
XX
PD 20-JAN-2000.
XX
PF 12-JUL-1999; 99WC-US15600.
XX
PR 13-JUL-1998; 98US-0092672.
PR 02-DEC-1998; 98US-0110608.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Thorpe PE, Ran S;
XX
DR WPI; 2000-182175/16.
DR N-PSDB; AA292215.
XX
XX New composition for killing tumour vascular endothelial cells for
PT treating solid tumours, comprises unconjugated anti-aminophospholipid
PT antibody -
XX
PS Disclosure; Page 223-224; 226pp; English.
XX
CC This sequence represents the human angiopoietin-2 (Ang-2) amino acid
CC sequence. Ang-2 is a naturally occurring angiogenesis antagonist when
CC vascular endothelial growth factor (VEGF) levels are low, and generally
CC counteracts the blood vessel maturation and stability mediated by Ang-1.
CC Ang-1 is a naturally occurring angiogenesis agonist, and is a maturation
CC or stabilisation factor, converting immature vessels to mature vessels.
CC Both Ang-1 and Ang-2 are useful in a therapeutic approach to the
CC treatment of vascularised tumours. The invention relates to a composition
CC comprising an anti-aminophospholipid antibody, or its antigen binding
CC region. The composition is used to kill tumour vasculature endothelial
CC cells. Aminophospholipids are stable and specific markers accessible on
CC the luminal surface of tumour blood vessels. Ang-1 or Ang-2 may be used
CC in the composition of the invention. The composition is used to treat
CC malignant or benign vascularised tumours in animals, especially large
CC tumours.
XX
SQ Sequence 496 AA;
Query Match 22.0%; Score 522; DB 21; Length 496;
Best Local Similarity 27.9%; Pred. No. 5.2e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;
QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ-LPKQFSRIEVEVFKEVQN 84
DB 26 MDSIGKKQYQVQHGSCSYTFJLPENDNCRSSSPYVSNAVGADAPLEYDCLSVQRLQVLEN 85
QY 85 LKE-----IVNSLKKSCQCKLQADGNGDFGRNGLLPSTGAPGEVGDN.... 128
DB 86 IMENNTQWLKLENYIQNMKKEMVEIQNAVN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSELKNAKEEINVLHGR---LEKL 161
DB 135 TAEQTRKLTDEVAQVLNQTTTRLEQLLEHSLSTNKLEKQILDQTSEINKLQDKNSFLEKK 194
QY 162 NLV-----NMNINIYVDSKVANLTF-----V 193
DB 195 VLAMEDKHTIQLSIKEEKDQLQVLSKQNSIEELEKKIVTATVNSVLSQKQCHDLMET 254

QY 184 VNSLDGKCKCPSQEQIQSRPVQHLYIKDCSDYYAIGKRSSEYRVTPTDPPKNSSFEVYCD 243
DB 255 VNNLLTMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 314
QY 244 YETMGSGWTVLQARLDGSTNFTWQDYKAGFGNLRREFWLGNDKIHLTKSKEMILRID 303
DB 315 MEAGGGWTIIQRREDGSDVDFORTWKEYKVGFGNPSGEYWLGNFEVSQLTNQORYVLKI 374
QY 304 LEDFNGVELYALYDQFYVANEFKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPDKDN 363
DB 375 LKCEWEGNEAYSLYEHFYLSSSEELNYRIHLKGLTGTAGKISSISQPGND----FSTKDGDN 430
QY 364 DRYPSGNCGLYYSSSGWFDACLSANLNGKYYHQYRGVR-NGIFWGTWPGVSEAHPGGYK 422
DB 431 DKCIC-KCSQMLTGGWFDACGPSNLNGMYYPQRONTNKFKNGIKWYYWKG-----SGY- 482
QY 423 SSFKEAKMIRPKHF 437
DB 483 -SLKATTMIRPADF 496
RESULT 10
AAU77944
ID AAU77944 standard; Protein; 496 AA.
XX
AC AAU77944;
XX
DT 02-JUL-2002 (first entry)
XX
DE Amino acid sequence for human angiopoietin-2.
XX
KW Human; immunoconjugate; anti-vascular endothelial growth factor antibody;
KW anti-VEGF antibody; monoclonal antibody 2C3 ATCC PTA 1595; VEGF receptor;
KW VEGFR2; KDR/Flk-1; VEGFR1; Flt-1; angiogenesis; macular degeneration;
KW ocular neovascular disease; cancer; vascularised solid tumour; AIDS;
KW metastatic tumour; endothelial cell proliferation; inflammatory disorder;
KW atherosclerosis; diabetic retinopathy; corneal graft rejection;
KW acquired immune deficiency syndrome; infection; restenosis; fungal ulcer;
KW sickle cell anaemia; endometriosis; angioipoetin-2.
XX
OS Homo sapiens.
XX
PN AU200179401-A.
XX
PD 06-DEC-2001.
XX
PF 12-OCT-2001; 2001AU-0079401.
XX
PR 28-APR-2000; 2000AU-0048049.
PR 12-OCT-2001; 2001AU-0079401.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Thorpe PE, Brekken RA;
XX
DR WPI; 2002-281368/33.
DR N-PSDB; ABK47716.
XX
XX Immunoconjugate compositions for treating cancer by inhibiting
PT angiogenesis and for delivering a diagnostic agent to tumour, comprises
PT anti-vascular endothelial growth factor antibody attached to a
PT biological agent -
XX
PS Disclosure; Page 4-6 (Sequence listing); 300pp; English.
XX
CC The present invention relates to antibody-based compositions comprising
CC an immunoconjugate such as anti-vascular endothelial growth factor
CC (VEGF) antibody (Ab) (or its antigen-binding fragment), attached to a
CC biological agent, where the Ab binds to the same epitope as the
CC monoclonal antibody (MAb) 2C3 ATCC PTA 1595, and significantly inhibits
CC VEGF binding to the VEGF receptor VEGFR2 (KDR/Flk-1) without inhibiting
CC VEGF binding to the VEGF receptor VEGFR1 (Flt-1). The compositions

Db 315 MEAGGGWTIIQRREDGSDVFQRTWKEYKVGFGNPSGEYKWLGNFVSQLTNQRYVLKIH 374

QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTTDPKDN 363

Db 375 LKDWEGNEAYSLEYEHFYLSEELNYRIHLKGLTGTAGKISSISQPGND----FSTKGDGN 430

QY 364 DRYPSGNCGLYYSSGWNFDACLSANLNGKYHHQYRGVR-NGIFGWTWPGVSEAHFPGGYK 422

Db 431 DKCIC-KCSQMLTGGWFWFDACGPSNLNGMYYPQRQNTNKFNGIKWYYWKG-----SGY- 482

QY 423 SSFKEAKMMIRPKHF 437

Db 483 -SLKATMMIRPADF 496

RESULT 12

ABP58064

ID ABP58064 standard; Protein: 496 AA.

XX

AC ABP58064;

XX

DT 07-MAR-2003 (first entry)

DE Human angiopoietin-2.

XX

KW Gene therapy; vector; hepatitis B virus; cardiovascular disease;

KW heart; cardiant; vasotropic; antiarrhythmic; antiarteriosclerotic;

KW human; angiopoietin-2.

XX

CS Homo sapiens.

XX

EN WC200287594-A1.

PC 07-NOV-2002.

XX

PF 30-APR-2002; 2002WO-US13644.

XX

PR 30-APR-2001; 2001US-287423P.

XX

PA (REGC) UNIV CALIFORNIA.

XX

PI Chien KR, Hoshijima M;

XX

DR WPI; 2003-111844/10.

XX

PT Novel non-viral vector comprises vesicular membrane with hepatitis B

PT envelope protein with cardiac targeting sequence, and nucleotide

PT sequence for gene therapy useful for treating, e.g., heart failure,

PT arrhythmia and atherosclerosis -

XX

PS Disclosure; Page 38-40; 53pp; English.

XX

CC The present sequence is the protein sequence of human

CC angiopoietin-2. The invention provides a non-viral vesicle vector

CC for the delivery of nucleic acid to various cardiac cell types.

CC The vesicle vector contains the hepatitis B virus envelope protein

CC in which at least part of the liver targeting sequence is deleted

CC and replaced with a specific cardiac cell targeting sequence. For

CC example, to selectively target vascular endothelial cells, peptides

CC including angiopoietin-2 are used. The vesicle vector can be

CC delivered intravenously or intra-arterially rather than by more

CC invasive methods such as direct cardiac injection. It can be used

CC to deliver gene products to replace or enhance expression of

CC proteins for treatment of heart failure, arrhythmia, reperfusion

CC injury, atherosclerosis, to promote angiogenesis, etc. The vesicles

CC are highly stable and can be produced in large quantities, making

CC them ideal for gene therapy.

XX

SQ Sequence 496 AA;

Query Match 22.0%; Score 522; DB 24; Length 496;

Best Local Similarity 27.9%; Pred. No. 5.2e-38;

Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVQN 84

Db 26 MDSIGKKQYQVQHGSCTYTLFLPEMDNCRSSSSPYVSNVQVQDAPLEYDDSVQRLQVLEN 85

QY 85 LKE-----IVNSLKKSQDCKLQADDNGDPGRNGLLLPSTGAPGEVGDN---- 128

Db 86 IMENNTQWLKLENYIQDNMKKEMVEIQQNAVQN-----QTAVMIEIGTNLLNQ 134

QY 129 -----RVRELESEV-----NKLSSSELKNVAKKEEINVLHGR---LEKL 161

Db 135 TAEQTRKLTQVQAQVLNQTTTRLEQLLEHSLSTNKLKQLDQTSINKLQDKNSFLEKK 194

QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183

Db 195 VLAMEDKHIIQJOSIKEEKDQLQVLVSKQNSIIEELEKKIVTATVNSVLQKQHQHDLMET 254

QY 184 VNSLDGKCKSPSQEQIQSRPVQHLIYKDCSDYYAIGKRSETYRVTPDKNSSFEVYCD 243

Db 255 VNNLLTMVSTNSAKCPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 314

QY 244 METYGGGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLITKSKEMILRID 303

Db 315 MEAGGGWTIIQRREDGSDVFQRTWKEYKVGFGNPSGEYKWLGNFVSQLTNQRYVLKIH 374

QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTTDPKDN 363

Db 375 LKDWEGNEAYSLEYEHFYLSEELNYRIHLKGLTGTAGKISSISQPGND---FSTKGDGN 430

QY 364 DRYPSGNCGLYYSSGWNFDACLSANLNGKYHHQYRGVR-NGIFGWTWPGVSEAHFPGGYK 422

Db 431 DKCIC-KCSQMLTGGWFWFDACGPSNLNGMYYPQRQNTNKFNGIKWYYWKG-----SGY- 482

QY 423 SSFKEAKMMIRPKHF 437

Db 483 -SLKATMMIRPADF 496

RESULT 13

ABG06771

ID ABG06771 standard; Protein: 572 AA.

XX

AC ABG06771;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #6762.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR N-PSDB; AAS70958.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

QY 162 NLV-----NMNMIENYVDSKVANLT?-----V 183
Db 732 VLAYEDKHIIQLOSIKEEKDQJQVLVSKQNSIIIEEJEKKIVATVANNVSVLQKQCHLMET 791
QY 184 VNSLDGKCKSKPCSQEQIQSRPVQHLLIYKDCSDYYAIGKRSSEYRVTPDPKNSSEYVYCD 243
Db 792 VNNLLTMVSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 651
QY 244 METMGGWTVLQARLDGSTNFTRWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMILRID 303
Db 852 MEAGGGWTIIQRREDGSDVDFQRTWKEYKVGFGNPSGEYWLGNFVSQLTNQCRYVLKI 911
QY 304 LEFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKFNHOLKFFTTEDKN 363
Db 912 LKDWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAGKISSISQPGND---FSTKCGDN 967
QY 364 DRYPSGNCGLYYSSGWGFDACLSANLNGKYYHQXYRGVR-NGIFWGTWPGVSEAHFGGYK 422
Db 968 DKCIC-KCSQMLTGGWGFDAACGSPSNLNGMYYPQRTNKNFNGIKWYYWKG-----SGY- 1019
QY 423 SSFKEAKMMIRPKHF 437
Db 1020 -SLKATTMMIRPADF 1033

RESULT 15
ID AAW47528 standard; Protein; 498 AA.

AC AAW47528;
XX
DT 09-SEP-1998 (first entry;
DE Amino acid sequence of chimeric TIE ligand IN1C2F (chimera 1).
XX
KW Chimeric TIE ligand IN1C2F; TIE-2 ligand; neovascularisation;
XX tumour; human.
OS Homo sapiens.
XX
PN W09805779-A1.
XX
PD 12-FEB-1998.
XX
PF 01-AUG-1997; 97WO-US13557.
XX
FR 25-OCT-1996; 96US-0740223.
FR 02-AUG-1996; 96US-0022999.
XX
PA (REGE-) REGENERON PHARM INC.
XX

PZ Davis S, Yancopoulos GD;
XX
DR WPI; 1998-145615/13.
DR N-PSDB; AAV18615.
XX

PT Modified human TIE-2 receptor ligand(s) - useful for promoting wound
PT healing
XX
FS Claim 20; Fig 24; 202pp; English.
XX

CC This is the amino acid sequence of the chimeric TIE ligand IN1C2F,
CC used in the method of the invention, involving the production
CC of TIE-2 ligands which promote healing. The nucleic acids, vectors
CC and host cells used in the method of the invention are useful for
CC the recombinant production of the ligands. The ligands, etc. are
CC useful for blocking blood vessel growth, promoting neovascularisation,
CC promoting the growth or differentiation of a cell expressing the TIE
CC receptor, blocking the growth or differentiation of a cell expressing
CC the TIE receptor and for attenuating or preventing tumour growth in
CC a human.
XX

SQ Sequence 498 AA;

Query Match 21.9%; Score 521.5; DB 19; Length 498;
Best Local Similarity 32.7%; Pred. No. 5.8e-38;
Matches 131; Conservative 61; Mismatches 138; Indels 71; Gaps 14;
QY 61 LPPLTIQLPKQFSRIE-----EVFKEVQNKLKEIVNSLKKSCQCKLQADENGCPGRN 112
Db 145 LTDVETQVLNQTSRLEIQLLNSLSTYKLEKQLLQQTNEILKIHE-----KN 191
QY 113 GLLLPSTGAPGEVDNRVRELESEVNLSSSELKNAKEEINVLHG-----RLEK-L 161
Db 192 SLL-----EHKILEMS---GKHKEELDTLKEEKENLQGLVTRCTYIIQELEKQL 237
QY 162 NLVNMNMIENYVDSKVANLTFVNSLSDGKCSK-----CPSQEIQSRPVQHLLIYKDCSDYY 217
Db 238 NRATTN--SVLQKQQJELMDTVHNLVNLCTKEGVLLKGGKREEKP-----FRDCAEVF 290
QY 218 AIGKRSSEYRVTPDPKNSSEFVYCDMETMGGWTVLQARLDGSTNFTRWQDYKAGFGN 277
Db 291 KSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGWTIIQRREDGSDVDFQRTWKEYKVGFGN 350
QY 278 LRREFWLGNDKIHLLTKSKEMILRIDLEDFNGVELYALYDQFYVANEFLKYRLHVGNYNG 337
Db 351 PSGEYWLGNFVSQLTNQCRYVLKIHLKDWEGNEAYSLEYEDFYLSSEELNYRIHLKGLTG 410
QY 338 TAGDALRFNKFNHDLKXFFTPTPKDNDKYPSPGNCGLYYSSGWGFDACLSANLNGKYYHQK 397
Db 411 TAGKISSISQPGND---FSTKCGDNDKCIC-KCSQMLTGGWGFDAACGSPSNLNGMYYPQR 465
QY 398 YRGVR-NGIFWGTWPGVSEAHFGGYKSSPKKAKMMIRPKHF 437
Db 466 QNTNKFNGIKWYYWKG-----SGY--SLKATTMMIRPADF 498

Search completed: November 5, 2003, 16:45:59
Job time : 100.059 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 16:48:06 ; Search time 49.315 Seconds
(without alignments)
1528.905 Million cell updates/sec

Title: US-09-902-563-2
Perfect score: 2378
Sequence: 1 MKLANWYWLSSAVLATYGF.....GYKSSFEXAKMMIRPKHKP 439

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2378	100.0	439	11	US-09-902-563-2 Sequence 2, Appli
2	2378	100.0	439	15	US-10-096-255-2 Sequence 2, Appli
3	1853.5	77.9	432	11	US-09-902-563-4 Sequence 4, Appli
4	1853.5	77.9	432	15	US-10-096-255-4 Sequence 4, Appli
5	525	22.1	496	15	US-10-263-677-9 Sequence 9, Appli
6	524	22.0	496	15	US-10-215-224-7 Sequence 7, Appli
7	522	22.0	480	15	US-10-225-060-8 Sequence 8, Appli
8	522	22.0	496	10	US-09-998-831-4 Sequence 4, Appli
9	522	22.0	496	10	US-09-897-306-14 Sequence 14, Appli
10	522	22.0	496	11	US-09-998-833-4 Sequence 4, Appli
11	522	22.0	496	12	US-10-136-819-13 Sequence 13, Appli
12	522	22.0	496	12	US-10-179-615-6 Sequence 6, Appli
13	522	22.0	496	12	US-10-179-820-6 Sequence 6, Appli
14	522	22.0	496	12	US-10-373-561-4 Sequence 4, Appli
15	522	22.0	496	14	US-10-179-744-6 Sequence 6, Appli

16	522	22.0	496	15	US-10-186-817-6 Sequence 6, Appli
17	522	22.0	496	15	US-10-215-224-8 Sequence 8, Appli
18	522	22.0	496	15	US-10-225-060-6 Sequence 6, Appli
19	522	22.0	496	15	US-10-263-677-8 Sequence 8, Appli
20	522	22.0	496	15	US-10-321-332-6 Sequence 6, Appli
21	521.5	21.9	498	15	US-10-225-060-20 Sequence 20, Appli
22	517.5	21.8	499	15	US-10-225-060-24 Sequence 24, Appli
23	510	21.4	496	15	US-10-225-060-15 Sequence 15, Appli
24	506	21.3	339	9	US-09-925-301-1082 Sequence 1082, Ap
25	505	21.2	491	10	US-09-897-306-1 Sequence 1, App-i
26	505	21.2	491	12	US-10-137-870-278 Sequence 278, App
27	505	21.2	491	12	US-10-140-018-278 Sequence 278, App
28	505	21.2	491	12	US-10-140-021-278 Sequence 278, App
29	505	21.2	491	12	US-10-140-274-278 Sequence 278, App
30	505	21.2	491	12	US-10-140-471-278 Sequence 278, App
31	505	21.2	491	12	US-10-140-807-278 Sequence 278, App
32	505	21.2	491	12	US-10-140-922-278 Sequence 278, App
33	505	21.2	491	12	US-10-140-924-278 Sequence 278, App
34	505	21.2	491	12	US-10-140-926-278 Sequence 278, App
35	505	21.2	491	12	US-10-141-698-278 Sequence 278, App
36	505	21.2	491	12	US-10-141-702-278 Sequence 278, App
37	505	21.2	491	12	US-10-141-704-278 Sequence 278, App
38	505	21.2	491	12	US-10-142-421-278 Sequence 278, App
39	505	21.2	491	12	US-10-142-432-278 Sequence 278, App
40	505	21.2	491	12	US-10-142-767-278 Sequence 278, App
41	505	21.2	491	12	US-10-143-033-278 Sequence 278, App
42	505	21.2	491	12	US-10-144-994-278 Sequence 278, App
43	505	21.2	491	12	US-10-145-628-278 Sequence 278, App
44	505	21.2	491	12	US-10-145-631-278 Sequence 278, App
45	505	21.2	491	12	US-10-145-633-278 Sequence 278, App

ALIGNMENTS

RESULT 1
US-09-902-563-2
; Sequence 2, Application US/09902563
; Publication No. US20030093654A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-37
; CURRENT APPLICATION NUMBER: US/09/902,563
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: US 09/442,143
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fgl2
US-09-902-563-2

Query Match	100.0%;	Score 2378;	DB 11;	Length 439;
Best Local Similarity	100.0%;	Pred. No. 1.4e-207;		
Matches 439;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKLANWYWLSSAVLATYGFLVANNETEIKDERAKVCPVRLSRGKCEEGECYQVS	60	
Db	1	MKLANWYWLSSAVLATYGFLVANNETEIKDERAKVCPVRLSRGKCEEGECYQVS	60	
QY	61	LPPLTIQLPKQFSRIEEVFKEVQNLIKVINSLKSCQDCKLQADDNGDPGRNGLLPSTG	120	
Db	61	LPPLTIQLPKQFSRIEEVFKEVQNLIKVINSLKSCQDCKLQADDNGDPGRNGLLPSTG	120	
QY	121	APGEVGDNRVRELESEVNKLSSSELKNKKEEINVHLGRLEKLVNMNINIYVDSKVANL	180	
Db	121	APGEVGDNRVRELESEVNKLSSSELKNKKEEINVHLGRLEKLVNMNINIYVDSKVANL	180	
QY	181	TFVNSLDGKCKPSEOIQSRPVCHLIYKDCSDYYAIGKRSSEYRVTCPKNSSEFEV	240	

Db 181 TFWVNSLDGKCKSPSQEQIQSRPVQHLYKDCSDYYAIGKASSETYRVTPDPKNSSEFV 240
Qy 241 YCDMETMGSGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMIL 300
Db 241 YCDMETMGSGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMIL 300
Qy 301 RIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPD 360
Db 301 RIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPD 360
Qy 361 KNDRYPSGNCGLYSSGWWFEDACLSANLNGKYYHQYRGVNRGIFWGTWPGVSEAHPGG 420
Db 361 KNDRYPSGNCGLYSSGWWFEDACLSANLNGKYYHQYRGVNRGIFWGTWPGVSEAHPGG 420
Qy 421 YKSSFKEAKMMIRPKHKFP 439
Db 421 YKSSFKEAKMMIRPKHKFP 439

RESULT 2
US-10-096-255-2
; Sequence 2, Application US/10096255
; Publication No. US20030103974A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-52
; CURRENT APPLICATION NUMBER: US/10/096,255
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fgl2
US-10-096-255-2

Query Match 100.0%; Score 2378; DB 15; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.4e-207;
Matches 439; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

Qy 1 MCLANWYWLSSAVLATYGFLLVANNETEEIKDERAKDVCVRLESRGKCEEAGECPYQVS 60
Db 1 MCLANWYWLSSAVLATYGFLLVANNETEEIKDERAKDVCVRLESRGKCEEAGECPYQVS 60

Qy 61 LPPLTIQLPKQFSRIEVEVFKEVQNKEIVNSLUKSCQDCKLQADNDGDPGRNGLLLPSTG 120
Db 61 LPPLTIQLPKQFSRIEVEVFKEVQNKEIVNSLUKSCQDCKLQADNDGDPGRNGLLLPSTG 120

Qy 121 APGEVGNRVRLESEVNVKLSSELKNAAKEEINVLHGRLEKLNLVNMNINENYVDSKVANL 180
Db 121 APGEVGNRVRLESEVNVKLSSELKNAAKEEINVLHGRLEKLNLVNMNINENYVDSKVANL 180

Qy 181 TFWVNSLDGKCKSPSQEQIQSRPVQHLYKDCSDYYAIGKASSETYRVTPDPKNSSEFV 240
Db 181 TFWVNSLDGKCKSPSQEQIQSRPVQHLYKDCSDYYAIGKASSETYRVTPDPKNSSEFV 240

Qy 241 YCDMETMGSGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMIL 300
Db 241 YCDMETYGGGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMIL 300

Qy 301 RIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPD 360
Db 301 RIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPD 360

Qy 361 KNDRYPSGNCGLYSSGWWFEDACLSANLNGKYYHQYRGVNRGIFWGTWPGVSEAHPGG 420
Db 361 KNDRYPSGNCGLYSSGWWFEDACLSANLNGKYYHQYRGVNRGIFWGTWPGVSEAHPGG 420

Qy 421 YKSSFKEAKMMIRPKHKFP 439
Db 421 YKSSFKEAKMMIRPKHKFP 439

RESULT 3
US-09-902-563-4
; Sequence 4, Application US/09902563
; Publication No. US20030093654A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-37
; CURRENT APPLICATION NUMBER: US/09/902,563
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: US 09/442,143
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Murine fgl2
US-09-902-563-4

Query Match 77.9%; Score 1853.5; DB 11; Length 432;
Best Local Similarity 77.7%; Pred. No. 6.7e-160;
Matches 341; Conservative 42; Mismatches 49; Indels 7; Gaps 4;

Qy 1 MCLANWYWLSSAVLATYGFLLVANNETEEIKDERAKDVCVRLESRGKCEEAGECPYQVS 60
Db 1 MRLPGWLWLSSAVLAACR-AVEEHNLTGLEDDASQAACPARLEGSGRC-EGSQCPFQJLT 58

Qy 61 LPPLTIQLPKQFSRIEVEVFKEVQNKEIVNSLUKSCQDCKLQADNDGDPGRNGLLLPSTG 120
Db 59 LPTLTITQLPRQLGSNEEVLKEVRTLKEAVDSLKSCQDCKLQADGHRDPGGNG---GNG 114

Qy 121 APGEVGNRVRLESEVNVKLSSELKNAAKEEINVLHGRLEKLNLVNMNINENYVDSKVANL 180
Db 115 AE-TAEDSRVQELSESQVNVKLSSELKNAKDQIQGLQGRLETLHLVNMNINENYVDNKVANL 173

Qy 181 TFWVNSLDGKCKSPSQEQIQSRPVQHLYKDCSDYYAIGKASSETYRVTPDPKNSSEFV 240
Db 174 TFWVNSLDGKCKSPSQEQEHMQSQPVQHLYKDCSDHYVVLGRRSSGAYRVTPDHRNSSFV 233

Qy 241 YCDMETMGSGWTVLQARLDGSTNFTRTWQDYKAGFGNLRREFWLGNDKIHLLTKSKEMIL 300
Db 234 YCDMETMGSGWTVLQARLDGSTNFTREWKDYKAGFGNLEREFWLGNDKIHLLTKSKEMIL 293

Qy 301 RIDLEDFNGVELYALYDQFYVANEFKYLRLHVGNYNGTAGDALRFNKHYNHDLKFFTTTPD 360
Db 294 RIDLEDFNGTLYALYDQFYVANEFKYLRLHIGNYNGTAGDALRFSRHYNHDLRFFTTTPD 353

Qy 361 KNDRYPSGNCGLYSSGWWFEDACLSANLNGKYYHQYRGVNRGIFWGTWPGVSEAHPGG 420
Db 354 RENDRYPSGNCGLYSSGWWFDSCLSANLNGKYYHQYKGVNRGIFWGTWPGINQAQPGG 413

Qy 421 YKSSFKEAKMMIRPKHKFP 439
Db 414 YKSSFKAQAKMIRPKHKFP 432

RESULT 4
US-10-096-255-4
; Sequence 4, Application US/10096255
; Publication No. US20030103974A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-52
; CURRENT APPLICATION NUMBER: US/10/096,255

QY 423 SSFKEAKMMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 9

US-09-897-306-14
; Sequence 14, Application US/09897306
; Patent No. US20020123054A1
; GENERAL INFORMATION:
; APPLICANT: Hillmat, Jennifer L.
; APPLICANT: Goigcne, Gina A.
; APPLICANT: Patterson, Chandra
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN ANGIOPOIETIN
; FILE REFERENCE: PC-0048 CIP
; CURRENT APPLICATION NUMBER: US/09/897,306
; CURRENT FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020123054A1 52257933
US-09-897-306-14

Query Match 22.0%; Score 522; DB 10; Length 496;
Best Local Similarity 27.9%; Pred. No. 1e-38;
Matches 136; Conservative 74; Mismatches 159; Indels 124; Gaps 15;
QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVQN 84
Db 26 MDSIGKKQYQVCHGSCSYTFLLPEMDNCRSSSPYVSNVQVORDAPLEYDDSDVQRLQVLEN 85
QY 85 LKE-----IVNSLKSCQDCKLQADNDGDPGRNGLLLPSTGAPGEVGDN---- 128
Db 86 IMENNTQWLKLENYIQNKYKKEVMEIQQNAVQN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSELKNAAKEEINVJLHGR---LEKL 161
Db 135 TAEQTRKLTQVEAQVNLQTTTRLEQLLEHSLSTNKLQKQILDQTSSEINKLQDNKSFLEKK 194
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
Db 195 VLAMEDKHIIQLQSIKEEKDQLQVLVSKQNSIIEELEKKIVTATVNSVLQKQOQHDLMET 254
QY 184 VNSLDGKCKSCPSQEQIQSRPVQHLYIKDCSDYYAIGKSSSETYRVTDPKNSSPFVYCD 243
Db 255 VNNLLTMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFGLRRREFWLGNDKIHLLTKSKEMILRID 303
Db 315 MEAGGGGTIIQRREDGSDVDFQRTWKEYKVGFNPSGGEYWLGNFVSQLTNQQRVVLKIH 374
QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTPDKDN 363
Db 375 LKOWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAQKISSISQPGND----FSTKQGEN 430
QY 364 DRYPSGNCGLYYSSGWWFDFACLSANLNGKYHQQYRGVR-NGIFWGTWPGVSEAHFPGYK 422
Db 431 DKCIC-KCSQMLTGGWFFDAGCPSNLNGMYYPQRQNTNKFNGIKWYWKG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 10

US-09-998-833-4

; Sequence 4, Application US/09998833
; Publication No. US20030082187A1
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING ANTIBODIES TO
; FILE REFERENCE: 4001.002200
; CURRENT APPLICATION NUMBER: US/09/998,833
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US/09/351,543
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-998-833-4

Query Match 22.0%; Score 522; DB 11; Length 496;
Best Local Similarity 27.9%; Pred. No. 1e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;
QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVQN 84
Db 26 MDSIGKKQYQVCHGSCSYTFLLPEMDNCRSSSPYVSNVQVORDAPLEYDDSDVQRLQVLEN 85
QY 85 LKE-----IVNSLKSCQDCKLQADNDGDPGRNGLLLPSTGAPGEVGDN---- 128
Db 86 IMENNTQWLKLENYIQNKYKKEVMEIQQNAVQN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSELKNAAKEEINVJLHGR---LEKL 161
Db 135 TAEQTRKLTQVEAQVNLQTTTRLEQLLEHSLSTNKLQKQILDQTSSEINKLQDNKSFLEKK 194
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
Db 195 VLAMEDKHIIQLQSIKEEKDQLQVLVSKQNSIIEELEKKIVTATVNSVLQKQOQHDLMET 254
QY 184 VNSLDGKCKSCPSQEQIQSRPVQHLYIKDCSDYYAIGKSSSETYRVTDPKNSSPFVYCD 243
Db 255 VNNLLTMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFGLRRREFWLGNDKIHLLTKSKEMILRID 303
Db 315 MEAGGGGTIIQRREDGSDVDFQRTWKEYKVGFNPSGGEYWLGNFVSQLTNQQRVVLKIH 374
QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFTPDKDN 363
Db 375 LKOWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAQKISSISQPGND----FSTKQGEN 430
QY 364 DRYPSGNCGLYYSSGWWFDFACLSANLNGKYHQQYRGVR-NGIFWGTWPGVSEAHFPGYK 422
Db 431 DKCIC-KCSQMLTGGWFFDAGCPSNLNGMYYPQRQNTNKFNGIKWYWKG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 11

US-10-136-819-13
; Sequence 13, Application US/10136819
; Publication No. US20030166593A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth
; APPLICANT: Hoshijima, Masahiko
; TITLE OF INVENTION: No. US20030166593A1-viral vesicle vector for cardiac specific g
; FILE REFERENCE: 6627-PAL198
; CURRENT APPLICATION NUMBER: US/10/136,819
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/287,423

QY 129 -----RVRELESEV-----NKLSELKNAKEEINVHGR---LEKL 161
DB 135 TAEQTRKLTDEVAQVLNQTTTRLEQLLEHSLSTNKLEKQILDQTSINKLQCKNSFLEKK 194
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
DB 195 VLAMEDKHIIQLQSIKEEKDQLQVLVSKQNSIIIEELEKKIVTATVNNSVLQKQHQHLMET 254
QY 184 VNSLDGKCKSPSQEQIQSRPVQHLLIYKDCSDYYAIGKRSSSETYRVTPDPKNSSEFEVYCD 243
DB 255 VNNLLTMSTNSAKOPTVAKEEQISFRDCAEVFKSGHTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLLTTSKEMILRID 303
DB 315 MEAGGGGTIIQRREDGSVDFTWKEYKVGFNPSGEYWLGNFEVFSQLTNQRYVLKIH 374
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLLTTSKEMILRID 303
DB 315 MEAGGGGTIIQRREDGSVDFTWKEYKVGFNPSGEYWLGNFEVFSQLTNQRYVLKIH 374
QY 304 LEDFNGVELYALYDQFYVANEFKYLRLHVGNNGTAGDALRPNKHYNHDLKFFTPDKDN 363
DB 375 LKOWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAGKISSISQPGND----FSTKDGDN 430
QY 364 DRYPSGNCGLYSSGWFWFACLSANLNGKYHQQYRGVR-NGIFWGTWPGVSEAHFPGYK 422
DB 431 DKCIC-KCSQMLTGGWWFDCGPNLNGMYYPQRTNKNFKNGIKWYYWKG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
DB 483 -SLKATMMIRPADF 496
RESULT 14
US-10-373-561-4
; Sequence 4, Application US/10373561
; Publication No. US20030175276A1
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002582
; CURRENT APPLICATION NUMBER: US/10/373,561
; PRIOR FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US/09/561,439
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-373-561-4

Query Match 22.0%; Score 522; DB 12; Length 496;
Best Local Similarity 27.9%; Pred. No. 1e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;
QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVQN 84
DB 26 MDSIGKKQYQVGHGSCSYTFLLPEDMENCNCRSSSPYVSNVAVQRDAPLEYDSDSVQRLQVLEN 85
QY 85 LKE-----IVNSLKKSCQDCKLQADNGDPGRNGLLLPSTGAPGEVGDN---- 128
DB 86 IMENNTQWLAKLENYICDNMKKEMVEIQQNAVQN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSELKNAKEEINVHGR---LEKL 161
DB 135 TAEQTRKLTDEVAQVLNQTTTRLEQLLEHSLSTNKLEKQILDQTSINKLQCKNSFLEKK 194
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
DB 195 VLAMEDKHIIQLQSIKEEKDQLQVLVSKQNSIIIEELEKKIVTATVNNSVLQKQHQHLMET 254

QY 184 VNSLDGKCKSPSQEQIQSRPVQHLLIYKDCSDYYAIGKRSSSETYRVTPDPKNSSEFEVYCD 243
DB 255 VNNLLTMSTNSAKOPTVAKEEQISFRDCAEVFKSGHTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLLTTSKEMILRID 303
DB 315 MEAGGGGTIIQRREDGSVDFTWKEYKVGFNPSGEYWLGNFEVFSQLTNQRYVLKIH 374
QY 304 LEDFNGVELYALYDQFYVANEFKYLRLHVGNNGTAGDALRPNKHYNHDLKFFTPDKDN 363
DB 375 LKOWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAGKISSISQPGND----FSTKDGDN 430
QY 364 DRYPSGNCGLYSSGWFWFACLSANLNGKYHQQYRGVR-NGIFWGTWPGVSEAHFPGYK 422
DB 431 DKCIC-KCSQMLTGGWWFDCGPNLNGMYYPQRTNKNFKNGIKWYYWKG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
DB 483 -SLKATMMIRPADF 496
RESULT 15
US-10-179-744-6
; Sequence 6, Application US/10179744
; Publication No. US20020173627A1
; GENERAL INFORMATION:
; APPLICANT: Davis, Samuel et al.
; TITLE OF INVENTION: TIE-2 LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG 330-F-PCT-US
; CURRENT APPLICATION NUMBER: US/10/179,744
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: US/08/817,318
; PRIOR FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
US-10-179-744-6

Query Match 22.0%; Score 522; DB 14; Length 496;
Best Local Similarity 27.9%; Pred. No. 1e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;
QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVQN 84
DB 26 MDSIGKKQYQVGHGSCSYTFLLPEDMENCNCRSSSPYVSNVAVQRDAPLEYDSDSVQRLQVLEN 85
QY 85 LKE-----IVNSLKKSCQDCKLQADNGDPGRNGLLLPSTGAPGEVGDN---- 128
DB 86 IMENNTQWLAKLENYICDNMKKEMVEIQQNAVQN-----QTAVMIEIGTNLLNQ 134
QY 129 -----RVRELESEV-----NKLSELKNAKEEINVHGR---LEKL 161
DB 135 TAEQTRKLTDEVAQVLNQTTTRLEQLLEHSLSTNKLEKQILDQTSINKLQCKNSFLEKK 194
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
DB 195 VLAMEDKHIIQLQSIKEEKDQLQVLVSKQNSIIIEELEKKIVTATVNNSVLQKQHQHLMET 254
QY 184 VNSLDGKCKSPSQEQIQSRPVQHLLIYKDCSDYYAIGKRSSSETYRVTPDPKNSSEFEVYCD 243
DB 255 VNNLLTMSTNSAKOPTVAKEEQISFRDCAEVFKSGHTNGIYTLTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLLTTSKEMILRID 303
DB 315 MEAGGGGTIIQRREDGSVDFTWKEYKVGFNPSGEYWLGNFEVFSQLTNQRYVLKIH 374
QY 304 LEDFNGVELYALYDQFYVANEFKYLRLHVGNNGTAGDALRPNKHYNHDLKFFTPDKDN 363

Db 375 LKDWEGNEAYSLEYEHFYLSSEELNYRIHLKGLTGTAGKISSISQPGND----FSTKDDGN 430
Qy 364 DRYPSGNCGLYYSSGWWFDACLSANLNGKYYHQKYRGVR-NQIFWGTWPGVSEAHFGGYK 422
Db 431 DKCIC-KCSQMJTGGWFDACGPNLNGMYYPQRQNTNKFNG:KYYYNKG-----SGY- 482
Qy 423 SSFKEAKMIRPKHF 437
Db 483 -SLKATTMIRPADF 496

Search completed: November 5, 2003, 16:56:34
Cob time : 52.315 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 16:44:23 ; Search time 30.9427 Seconds
(without alignments)
600.285 Million cell updates/sec

Title: US-09-902-563-2
Perfect score: 2378
Sequence: 1 MGLANWYWLSSAVLATYGF.....GYKSSFKKAKMIRPKHKFP 439

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/pcfus_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2378	100.0	439	US-09-442-143A-2	Sequence 2, Appli
2	1853.5	77.9	432	US-09-442-143A-4	Sequence 4, Appli
3	524	22.0	314	US-08-525-505A-2	Sequence 2, Appli
4	524	22.0	496	US-09-202-491-7	Sequence 7, Appli
5	522	22.0	480	US-08-740-223A-8	Sequence 8, Appli
6	522	22.0	480	US-09-709-188-8	Sequence 8, Appli
7	522	22.0	496	US-08-373-579-6	Sequence 6, Appli
8	522	22.0	496	US-08-418-595-6	Sequence 6, Appli
9	522	22.0	496	US-08-665-926-6	Sequence 6, Appli
10	522	22.0	496	US-09-162-437-6	Sequence 6, Appli
11	522	22.0	496	US-08-740-223A-6	Sequence 6, Appli
12	522	22.0	496	US-09-351-457-4	Sequence 4, Appli
13	522	22.0	496	US-09-561-500-4	Sequence 4, Appli
14	522	22.0	496	US-09-561-108-4	Sequence 4, Appli
15	522	22.0	496	US-09-351-543-4	Sequence 4, Appli
16	522	22.0	496	US-09-561-526-4	Sequence 4, Appli
17	522	22.0	496	US-09-202-491-8	Sequence 8, Appli
18	522	22.0	496	US-08-817-318-6	Sequence 6, Appli
19	522	22.0	496	US-09-709-188-6	Sequence 6, Appli
20	522	22.0	496	US-09-561-499-4	Sequence 4, Appli
21	521.5	21.9	498	US-08-740-223A-20	Sequence 20, Appl
22	521.5	21.9	498	US-09-709-188-20	Sequence 20, Appl
23	517.5	21.8	499	US-08-740-223A-24	Sequence 24, Appl
24	517.5	21.8	499	US-09-709-188-24	Sequence 24, Appl
25	510	21.4	496	US-08-740-223A-15	Sequence 15, Appl
26	510	21.4	496	US-09-709-188-15	Sequence 15, Appl
27	506.5	21.3	286	US-08-960-507-20	Sequence 20, Appl

28	506.5	21.3	286	4	US-09-136-801-20	Sequence 20, Appl
29	506.5	21.3	286	4	US-09-202-088A-20	Sequence 20, Appl
30	506	21.3	312	1	US-08-525-505A-4	Sequence 4, Appli
31	505	21.2	491	2	US-08-933-821-4	Sequence 4, Appli
32	505	21.2	491	3	US-08-960-507-4	Sequence 4, Appli
33	505	21.2	491	4	US-09-136-828-4	Sequence 4, Appli
34	505	21.2	491	4	US-09-332-928A-4	Sequence 4, Appli
35	505	21.2	491	4	US-09-136-801-4	Sequence 4, Appli
36	505	21.2	491	4	US-09-332-929-4	Sequence 4, Appli
37	505	21.2	491	4	US-09-333-075-4	Sequence 4, Appli
38	505	21.2	491	4	US-09-202-088A-4	Sequence 4, Appli
39	505	21.2	491	4	US-09-333-077-4	Sequence 4, Appli
40	500	21.0	491	4	US-09-658-644-2	Sequence 2, Appli
41	498	20.9	496	3	US-08-740-223A-16	Sequence 16, Appl
42	498	20.9	496	4	US-09-709-188-16	Sequence 16, Appl
43	489	20.6	346	3	US-08-960-507-19	Sequence 19, Appl
44	489	20.6	346	4	US-09-136-801-19	Sequence 19, Appl
45	489	20.6	346	4	US-09-202-088A-19	Sequence 19, App.

ALIGNMENTS

RESULT 1
US-09-442-143A-2
; Sequence 2, Application US/09442143A
; Patent No. 6403089
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation:
; FILE REFERENCE: 9579-14
; CURRENT APPLICATION NUMBER: US/09/442,143A
; CURRENT FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fg12
US-09-442-143A-2

Query Match:		100.0%;	Score 2378;	DB 4;	Length 439;
Best Local Similarity		100.0%;	Pred. No. 8e-211;		
Matches 439;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGLANWYWLSSAVLATYGLVANNETEEIKDERAKDVCVRLESRGKCEEAGECPYQVS	60		
Db	1	MGLANWYWLSSAVLATYGLVANNETEEIKDERAKDVCVRLESRGKCEEAGECPYQVS	60		
Qy	61	LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQDCKLQADDNGDPGRNGLLLPSTG	120		
Db	61	LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQDCKLQADDNGDPGRNGLLLPSTG	120		
Qy	121	APGEVGNRVRELESEVKNLSSELKNAKEEINVHGRLEKLNVMNNIENYVDSKVANL	180		
Db	121	APGEVGNRVRELESEVKNLSSELKNAKEEINVHGRLEKLNVMNNIENYVDSKVANL	180		
Qy	181	TFVNSLDGKCKCPSEQIQSRPVQHLYKDCSDYYAIGKRSSETYRVTPDPKNSSEFV	240		
Db	181	TFVNSLDGKCKCPSEQIQSRPVQHLYKDCSDYYAIGKRSSETYRVTPDPKNSSEFV	240		
Qy	241	YCDMETMGGWTVLQARLDGSTNFTRTWQYKAGFGNLRREFWJGNDKIHLTKSKEMIL	300		
Db	241	YCDMETMGGWTVLQARLDGSTNFTRTWQYKAGFGNLRREFWJGNDKIHLTKSKEMIL	300		
Qy	301	RIDLEDFNGLVYALYDQFYVANEFKYRLHVGNYNGTAGDALRFNKHYNHDLJKFFTTPD	360		
Db	301	RIDLEDFNGLVYALYDQFYVANEFKYRLHVGNYNGTAGDALRFNKHYNHDLJKFFTTPD	360		

Db 290 W-----RGWYSLKSVVMKIRPSDFIP 311

RESULT 4

US-09-202-491-7
; Sequence 7, Application US/09202491
; Patent No. 6432667
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG330-K
; CURRENT APPLICATION NUMBER: US/09/202,491
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: PCT/US97/10728
; EARLIER FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: 60/022,999
; EARLIER FILING DATE: 1996-08-02
; EARLIER APPLICATION NUMBER: 60/021,097
; EARLIER FILING DATE: 1996-07-02
; EARLIER APPLICATION NUMBER: 08/665,926
; EARLIER FILING DATE: 1996-06-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-202-491-7

Query Match 22.0%; Score 524; DB 4; Length 496;
Best Local Similarity 32.9%; Pred. No. 7.8e-40;
Matches 127; Conservative 65; Mismatches 154; Indels 40; Gaps 9;

QY 61 LPPLTIQLPKQFSRIEVEVPKEVQNLKEIVNSLKSCQDCKLQADDNGDPGRNGLLLPSTG 120
Db 142 LTDVEAQVNLQTFAL-----ELQLQHSISLNKLEKQILDQTSSEINKLNKNSFL----- 191

QY 121 APGEVGNRVRELESEVKNKLSSELKNAXEEINVLHGR-----EKLKLVNMNIENY 172
Db 192 -----EQVLVDXEGHSEQLQSMKEQDELQVLVSKQSSVIDELEKKLVATVNN--SL 243

QY 173 VESKVANLTFVNSLDGKCKSPSQEQIQSRPVQHLIYKDCSDYYAIGKRSSEYRVTPTD 232
Db 244 LCKCOHDLMEFVNSLTLTWYSSPNSKSSVAIRKEEQTFRCABIFKSLTSGIYTLTPP 303

QY 233 PKNSSFEVYCDMEFMGGGWTVLQARLDGSTNFTWQDYKAGFGNLRREFWLGNDKIHLL 292
Db 304 NSTEEIKAYCDMEVGGGGWTVIOHREDGSDVDFQRTWKEYKEGFGNPLGEYWLGNFVSQ 363

QY 293 TKSSEMILRIDLEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYND 352
Db 364 TGOHRYVLKIQKDWEGNEAHSLYDHFYLAGESNYRIHLTGLTGTAAKISSISQPGSD- 422

QY 353 LKFFTTPDKNDRYPSGNCGLYSSGWNWFDACLSANLNGKYYHQYRGVR-NGIFWGTWP 411
Db 423 ---FSTKCSNDKIC-C-KCSQMLSGGWNWFDACGPSNLNGQYYPQKQNTNKNFGIKWYWK 478

QY 412 GVSEAHPGGYKSSFKAKMVIKXHF 437
Db 479 G-----SGY--SLKATTMVIRPADF 496

RESULT 5

US-08-740-223A-8
; Sequence 8, Application US/08740223A
; Patent No. 6265564
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular
; TITLE OF INVENTION: Inter cellular Signalling Molecule
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.

; STREET: 777 Old Saw Mill Road
; City: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,223A
; FILING DATE: 25-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/022/999
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 333
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Mature TL2 protein
; LOCATION: 1..480
; OTHER INFORMATION:
US-08-740-223A-8

Query Match 22.3%; Score 522; DB 3; Length 480;
Best Local Similarity 27.9%; Pred. No. 1.1e-39;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE--EAGECPYQVSLPPL-----TIQ--LPKQFSRIEVEVPKEVQN 84
Db 10 MDSIGKKQYQVQHSGCSYTFLLPEMDNCRSSSPYVNAVCRDAPLEYDDSVORLQVLEN 69

QY 85 LKE-----IVNSLKSCQDCKLQADDNGDPGRNGLLLPSTGAPGEVGN---- 128
Db 70 IMENNTQWLKLENYIQDNMKKEMVEIQCNVQN-----QTAVMIEIGTNLLNQ 118

QY 129 -----RVRELESEV-----NKLSELKNAXEEINVLHGR----LEKL 161
Db 119 TAEQTRKLTDEVAQVNLQTRLEQLJLEHSLSTNKLKQILDQTSSEINKLQDKNSFLEKK 178

QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
Db 179 VLAMEDKHIIQLQSIKEKDQLQVLVSKQNSIIEELEKKIVTATVNSVLQKQKHDLMET 238

QY 184 VNSLDGKCKSPSQEQIQSRPVQHLIYKDCSDYYAIGKRSSEYRVTPTDPKNSSEFVYCD 243
Db 239 VNNLLTMMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCD 298

QY 244 METMGGGWTVLQARLDGSTNFTWQDYKAGFGNLRREFWLGNDKIHLLTKSEMILRID 303
Db 299 MEAGGGGWTVIQRREDGSDVDFQRTWKEYKVGFGNPSGEYWLGNFVSQLTNQRYVLKIH 358

QY 304 LEDFNGVELYALYDQFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNDLKFFTPDKDN 363
Db 359 LKDWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTGTAKEISSISQPGND----FSTKDDGN 414

QY 364 DRYPSGNCGLYSSGWNWFDACLSANLNGKYYHQYRGVR-NGIFWGTWPGVSEAHPGGYK 422
Db 415 DKCIC-KCSQMLTGGWNWFDACGPSNLNGMYYPQKQNTNKNFGIKWYWK-----SGY- 466

Best Local Similarity 27.9%; Pred. No. 1.2e-39;		
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;		
Qy	43	LESRGKE---EAGECPYQVSJPPD-----TIQ--LPKQFSRIEEVFKEVQN 84
Db	26	MDSIGKKQYQVQHGSCTYTFLLPENDCRSSSPYVSNVQORDAPLEYDCSVQRJQVLEN 85
Qy	85	LKE-----IVNSLKKSCQDCKLQADDNGCPGRNSGULLPSTGAPGEVGN---- 128
Db	86	IYENNTQWLMKLENYIQDNMKKEMVEIQNAVQN-----QTAVMIEIGTNLLNO 134
Qy	129	-----RVRELESEV-----NKLSELKNAAKEEINVLRGR---LEKL 161
Db	135	TAEQTRKLTQVEAQVLNQTTRLEQLJLHSLSTNKLKQCILQQTSEINKLQDKNSFLEKK 194
Qy	162	NLV-----NANIENYVDSKVANLTF-----V 183
Db	195	VLAEDKHIIQLQSIKEEKQQLQVLVSKQNSIIIELEKKVTAATVNSVLQKQCHDLMET 254
Qy	184	VNSLDGKCKCPQSOEQIQSRPVQHLYKDCSDYYAIGKRSETYRVTPDPKNSSEFYCD 243
Db	255	VNNLLTMSTNSAKDPTVAKEEQISPRDCAEVFKSGHTTNGIYTLTPFNSTEEIKAYCD 314
Qy	244	METMGSGWTVLQARLDGSTNFTWQYKAGFGNLRREFNLGNDKIHLLTKSKEMILRID 303
Db	315	MEAGGGWTTIQRREDGSVDFQRTWKEYKVGFGNFSGEYNLGNFVSQLTNQORYVLKIH 374
Qy	304	LEDFNGVEIYALYDQFYVANEELKYRLHVGNYNGTAGDALRPNKHYNHDLKFFETTRDKN 363
Db	375	LKWEGNEAYSLYEHFYLSSSEELNYRIHLKGLTGTAGKISSISQPGND----PSTKDCEN 430
Qy	364	DRYPNGNGLYYSSGWWFDACLSANLNGKYYHQYRGVR-NGLFWGTWPGVSEAHPSGYK 422
Db	431	DKCIC-KCSQMLTGGWWFDACGPNLNGYXYYPQRCNTNKNFNGIKHYWKG-----SGY- 482
Qy	423	SSFKEAKMMIRPKHF 437
Db	483	-SLKATTMMIRPADF 496

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 16:39:17 ; Search time 0.92511 Seconds
(without alignments)
1559.306 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPSGNCGLYYSSG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 36168692 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	88	100.0	432	2 I56934	fibrinogen-like pr
2	88	100.0	432	2 A27447	cytotoxic T-lympho
3	88	100.0	439	2 I37391	fibrinogen-like pr
4	46	52.3	2403	2 A59386	sanko - human
5	46	52.3	3461	2 S58870	reelin precursor -
6	43	48.9	225	2 C82893	hypothetical prote
7	43	48.9	284	2 I49707	germ cell specific
8	43	48.9	339	2 T24007	hypothetical prote
9	43	48.9	456	2 C86624	hypothetical prote
10	43	48.9	456	2 H72000	hypothetical prote
11	42.5	48.3	696	2 G71829	probable outer mem
12	41	46.6	273	2 T16246	hypothetical prote
13	41	46.6	368	2 A81289	hypothetical prote
14	41	46.6	385	2 T26404	hypothetical prote
15	41	46.6	437	2 A72513	hypothetical prote
16	41	46.6	517	1 ERADA7	early E2A DNA-bind
17	41	46.6	591	1 WMBPQ2	gene P2 protein -
18	41	46.6	680	2 H70347	outer membrane pro
19	41	46.6	798	2 T00131	xylan 1,4-beta-xylo
20	40	45.5	92	2 A72242	ferredoxin - Therm
21	40	45.5	100	2 JN0734	hypothetical 11.49
22	40	45.5	165	2 D86894	hypothetical prote
23	40	45.5	245	2 G64210	uracil DNA glycosyl
24	40	45.5	286	2 A33546	actin-capping prot
25	40	45.5	294	2 AF3043	calcium-binding pr
26	40	45.5	305	2 F98242	regucalcin (AB0379
27	40	45.5	455	2 E64454	proline-tRNA ligas
28	40	45.5	474	2 ERAD41	early E2A DNA-bind
29	40	45.5	484	2 ERAD12	early E2A DNA-bind

30 40 45.5 490 2 T44576 L-2,4-diaminobutyr
31 40 45.5 521 2 D87413 hypothetical prote
32 40 45.5 776 2 S28258 androgen-regulated
33 40 45.5 1151 2 T18297 zinc-finger protei
34 40 45.5 1211 2 S54500 alpha,alpha-trehal
35 39 44.3 117 2 I28195 Ig heavy chain V r
36 39 44.3 118 2 S25001 ubiquitin/ribosoma
37 39 44.3 155 2 JH0226 ubiquitin / riboso
38 39 44.3 155 2 JH0227 ubiquitin / riboso
39 39 44.3 155 2 JS0657 ubiquitin / riboso
40 39 44.3 156 1 UQTO7A ubiquitin / riboso
41 39 44.3 156 2 S25305 ubiquitin / riboso
42 39 44.3 156 2 T52334 ubiquitin extensio
43 39 44.3 156 2 T52335 ubiquitin extensio
44 39 44.3 157 2 D36571 ubiquitin / riboso
45 39 44.3 157 2 C36571 ubiquitin / riboso

ALIGNMENTS

RESULT 1
I56934
fibrinogen-like protein - mouse
C/Species: Mus sp. (mouse)
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 03-Dec-1999
C/Accession: I56934
R/Farr, R.L.; Fung, L.; Reneker, J.; Myers-Mason, N.; Leibowitz, J.L.; Levy, G.
J. Virol. 69, 5033-5038, 1995
A/Title: Association of mouse fibrinogen-like protein with murine hepatitis virus-indu
A/Reference number: I56934; MJID:95333285; PMID:7609073
A/Accession: I56934
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-432 <RES>
A/Cross-references: GB:S78773; NID:gl042169; PIDN:AAB34823.1; PID:gl042170
C/Genetics:
A/Gene: musfiblp
C/Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
F;203-428/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 100.0%; Score 88; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. NO. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 DRYPSGNCGLYYSSG 15
|||||
Db 357 DRYPSGNCGLYYSSG 371

RESULT 2
A27447
cytotoxic T-lymphocyte-specific protein precursor (clone pT49) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 15-Dec-1998 #sequence_revision 15-Dec-1988 #text_change 03-Dec-1999
C/Accession: A27447
R/Koyama, T.; Hall, L.R.; Haser, W.G.; Tonegawa, S.; Saito, H.
Proc. Natl. Acad. Sci. U.S.A. 84, 1609-1613, 1987
A/Title: Structure of a cytotoxic T-lymphocyte-specific gene shows a strong homology t
A/Reference number: A27447; MUID:87175527; PMID:3550794
A/Accession: A27447
A/Molecule type: mRNA
A/Residues: 1-432 <KOY>
A/Cross-references: GB:M16238; NID:gl93304; PIDN:AAA37624.1; PID:g387156
C/Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
F;203-428/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 100.0%; Score 88; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. NO. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DRYPSGNCGLYYSSG 15
|||||

Ds 357 DRYPSGNCGLYSSG 371

RESULT 3

I37391
fibrinogen-like protein expressed in T lymphocytes (p149) - human
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000
C:Accession: I37391; S47273
R:Ruegg, C.; Pytela, R.
Gene 160, 257-262, 1995

A:Title: Sequence of a human transcript expressed in T-lymphocytes and encoding a fibrin
A:Reference number: I37391; MUID:95369700; PMID:7642106

A:Accession: I37391

A>Status: preliminary; translated from GB/EMBL/DBEJ

A:Molecule type: mRNA

A:Residues: 1-439 <RES>

A:Cross-references: EXBL:Z36531; NID:G535184; PID:CAA85298.1; PID:G535185

A>Note: submitted to the EMBL Data Library, August 1994

C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology

F:210-435/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 100.0%; Score 88; DB 2; Length 439;

Best Local Similarity 100.0%; Pred. No. 2.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRYPSGNCGLYSSG 15

|||||

Ds 364 DRYPSGNCGLYSSG 378

RESULT 4

A59386

sanko - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-2001 #sequence_revision 31-Dec-2001 #text_change 31-Dec-2001

C:Accession: A59386

R:sanko, S.

submitted to the Protein Sequence Database, March 2001

A:Reference number: A59386

A:Accession: A59386

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-2403 <SAN>

Query Match 52.3%; Score 46; DB 2; Length 2403;

Best Local Similarity 64.3%; Pred. No. 54;

Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 4 PSNGCG--LYYSSG 15

|||||

Ds 1752 PSSNCGGFLFYASG 1765

RESULT 5

S58870

reelin precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 11-Jan-2000

C:Accession: S58870; S71844; I49297

R:D'Arcangelo, G.; Miao, G.G.; Chen, S.C.; Soares, H.D.; Morgan, J.J.; Curran, T.

Nature 374, 719-723, 1995

A:Title: A protein related to extracellular matrix proteins deleted in the mouse mutant

A:Reference number: I49297; MUID:95231649; PMID:7715726

A:Accession: S58870

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3461 <DAR>

A:Cross-references: EMBL:U24703; NID:G902486; PID:G902487

R:D'Arcangelo, G.

submitted to the EMBL Data Library, April 1995

A:Reference number: S71844

A:Accession: S71844

A:Molecule type: mRNA
A:Residues: 1-215, 'T', 217-1905, 'S', 1907-3355, 'V', 3357-3391, 'N', 3393-3461 <DA2>
A:Cross-references: EMBL:U24703; NID:G902486; PID:G902487
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-3461/Product: reelin #status predicted <MAT>
F:1769-1795/Domain: EGF homology <EGF>

Query Match 52.3%; Score 46; DB 2; Length 3461;

Best Local Similarity 71.4%; Pred. No. 76;

Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 3 YPSGNCGLY--YSS 14

|||||

Ds 1974 YPGNIGLYCPYSS 1987

RESULT 6

C82893

hypothetical protein UU425 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: C82893

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Casselli, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A:Reference number: A82870

A:Accession: C82893

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-225 <GLA>

A:Cross-references: GB:AE002139; GB:AF222894; NID:G6899405; PID:AAF30837.1; GSPDB:GNO

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: UU425

A:Genetic code: SGC3

Query Match 48.9%; Score 43; DB 2; Length 225;

Best Local Similarity 70.0%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YPSGNCGLY 12

|||||

Ds 87 YFNGNAELYY 96

RESULT 7

I49707

germ cell specific gene 3 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 19-May-2000

C:Accession: I49707

R:Tanaka, H.; Yoshimura, Y.; Nishina, Y.; Nozaki, M.; Nojima, H.; Nishimura, Y.

FEBS Lett. 355, 4-10, 1994

A:Title: Isolation and characterization of cDNA clones specifically expressed in testis

A:Reference number: I49707; MUID:95046372; PMID:7957958

A:Accession: I49707

A>Status: preliminary; translated from GB/EMBL/DBEJ

A:Molecule type: mRNA

A:Residues: 1-284 <RES>

A:Cross-references: GB:D38545; NID:G603492; PID:G603493

C:Superfamily: actin-capping protein alpha chain

Query Match 48.9%; Score 43; DB 2; Length 284;

Best Local Similarity 60.0%; Pred. No. 22;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DRYPSGNCGL 10

|||||

Ds 131 DRYPNGNCNV 140

RESULT 8

```
T24007
hypothetical protein R07B5.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T24007
R:Kelly, P.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19829
A:Accession: T24007
A:Status: preliminary; translated from GB/EMBL/CDDBJ
A:Molecule type: DNA
A:Residues: 1-339 <WIL>
A:Cross-references: EMBL:Z72512; PIDN:CAA96667.1; GSPDB:GNCC0023; CESP:R07B5.7
A:Experimental source: Clone R07B5
C:Genetics:
A:Gene: CESP:R07B5.7
A:Map position: 5
A:Introns: 31/2; 93/1; 124/3; 194/3; 247/3; 267/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C33A12.9b

Query Match 48.9%; Score 43; DB 2; Length 339;
Best Local Similarity 53.8%; Pred. No. 26;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYSSG 15
   |||||
Db 232 YPFGSVSVYSSG 244

RESULT 9
C86624
hypothetical protein CPJ1070 [imported] - Chlamydophila pneumoniae (strain J139)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: C86624
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ishiguro, M.
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J139.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: C86624
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-456 <STO>
A:Cross-references: GB:BA000008; NID:g8979443; PIDN:BA09277.1; GSPDB:GNCC0142
A:Experimental source: strain J139
C:Genetics:
A:Gene: CPJ1070

Query Match 48.9%; Score 43; DB 2; Length 456;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 YPSGNCGLY 12
   |||||
Db 3 FPCGNCNCY 12

RESULT 10
H72000
hypothetical protein CP0780 [imported] - Chlamydophila pneumoniae (strains CWL029 and AR)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: H72000; A81538
R:Kallan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Clinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: H72000
A:Molecule type: DNA
A:Residues: 1-456 <ARN>
A:Cross-references: GB:AE001687; GB:AE001363; NID:g4377398; PIDN:AAD19207.1; PID:g437740
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
```

```
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Saizber
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81538
A:Molecule type: DNA
A:Residues: 1-456 <REA>
A:Cross-references: GB:AE002238; GB:AE002161; NID:g7189693; PIDN:AAF38579.1; PID:g7189
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CPn1070; CP0780
```

```
Query Match 48.9%; Score 43; DB 2; Length 456;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 YPSGNCGLY 12
   |||||
Db 3 FPCGNCNCY 12
```

RESULT 11

```
G71829
probable outer membrane protein - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: G71829; G71960
R:Alm, R.A.; Ling, L.S.L.; McI, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, S.F.
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pa
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71829
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-696 <ARN>
A:Cross-references: GB:AE001549; GB:AE001439; NID:g4155858; PIDN:AAD06834.1; PID:g4155
A:Experimental source: strain J99
A:Accession: G71960
A:Molecule type: DNA
A:Residues: 1-696 <AR2>
A:Cross-references: GB:AE001459; GB:AE001439; NID:g4154723; PIDN:AAD05786.1; PID:g4154
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0212
```

```
Query Match 48.3%; Score 42.5; DB 2; Length 696;
Best Local Similarity 69.2%; Pred. No. 61;
Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 3 YPSGNCGLYSSG 15
   |||||
Db 231 YPCGN-GNYSSG 242
```

RESULT 12

```
T16246
hypothetical protein F35A5.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C:Accession: T16246
R:Leimbach, D.
submitted to the EMBL Data Library, January 1996
A:Description: The sequence of C. elegans cosmid F35A5.
A:Reference number: Z18485
A:Accession: T16246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-273 <LEI>
A:Cross-references: EMBL:U46675; NID:g1166613; PID:g1166614; PIDN:AAB52645.1; GSPDB:GN
A:Experimental source: strain Bristol N2; clone F35A5
C:Genetics:
```

A;Gene: CESP:F35A5.4

A;Map position: X

A;Introns: 18/3; 62/3; 81/3; 100/3; 140/3; 226/2

Query Match 46.6%; Score 41; DB 2; Length 273;

Best Local Similarity 66.7%; Pred. No. 44;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PSNCGLYYSSG 15

|||||

Db 50 PSNCGCGSCSG 61

RESULT 13

A81289

hypothetical protein Cj1433c (imported) - Campylobacter jejuni (strain NCTC 11168)

C;Species: Campylobacter jejuni

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C;Accession: A81289

R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.W.; Churcher, C.; Bashan, D.; Chilling

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre

Nature 403, 665-668, 2000

A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A;Reference number: A81250; MUID:20150912; PMID:10688204

A;Accession: A81289

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-368 <PAR>

A;Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73857.1; PID:G696886

A;Experimental source: serotype O2, strain NCTC 11168

C;Genetics:

A;Gene: C51433c

C;Superfamily: Campylobacter jejuni hypothetical protein Cj1433c

Query Match 46.6%; Score 41; DB 2; Length 368;

Best Local Similarity 70.0%; Pred. No. 59;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSNCGLYYYS 13

|||||

Db 265 PSNCNDYHS 274

RESULT 14

T26404

hypothetical protein Y105C5B.bb - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C;Accession: T26404

R;McMurray, A.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z20208

A;Accession: T26404

A;Status: preliminary; translated from GB/EMBL/DBEJ

A;Molecule type: DNA

A;Residues: 1-385 <WIL>

A;Cross-references: EMBL:AL110479; PIDN:CAB54376.1; CESP:Y105C5B.bb

A;Experimental source: clone Y105C5B

C;Genetics:

A;Gene: CESP:Y105C5B.bb

A;Introns: 58/1; 112/1; 270/2

C;Superfamily: glutamate-ammonia ligase

Query Match 46.6%; Score 41; DB 2; Length 385;

Best Local Similarity 53.3%; Pred. No. 61;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DRYPSCGLYYSSG 15

|||||

Db 341 DRPSSNCDFYTVTG 355

RESULT 15

A72513

hypothetical protein APE2078 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C;Accession: A72513

R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tak

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aerof

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: A72513

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-437 <KAW>

A;Cross-references: DDBJ:AP000063; NID:G5105654; PIDN:BAA81089.1; PID:G5105777

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE2078

C;Superfamily: conserved hypothetical protein MTH1394

Query Match 46.6%; Score 41; DB 2; Length 437;

Best Local Similarity 57.1%; Pred. No. 63;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 RYPSGNCGLYYSSG 15

|||||

Db 112 RFYEGEAGLYLSSG 125

Search completed: November 5, 2003, 16:48:15

Job time : 2.92511 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 15:59:02 ; Search time 0.560793 Seconds
(without alignments)
1067.506 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPSGNCGLYYSSG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	88	100.0	432	1 FGL2_MOUSE	P12804 mus musculus
2	88	100.0	439	1 FGL2_HUMAN	Q14314 homo sapien
3	46	52.3	3209	1 RELN_CHICK	Q93574 gallus gall
4	46	52.3	3460	1 RELN_HUMAN	P78509 homo sapien
5	46	52.3	3461	1 RELN_MOUSE	Q60841 mus musculus
6	46	52.3	3462	1 RELN_RAT	P58751 rattus norv
7	45	51.1	210	1 BCAS3_HUMAN	Q9N931 homo sapien
8	43	48.9	299	1 CAZ3_HUMAN	Q96KX2 homo sapien
9	43	48.9	299	1 CAZ3_MOUSE	P70190 mus musculus
10	43	48.9	299	1 CAZ3_RAT	Q9WJY6 rattus norv
11	41	46.6	121	1 CHA9_LYMDI	P50603 lymantria d
12	41	46.6	437	1 YK78_AERPE	Q9YA60 aeropyrum p
13	41	46.6	498	1 CTAC_HUMAN	Q96NM4 homo sapien
14	41	46.6	517	1 DNB2_ADE07	P04497 human adeno
15	41	46.6	590	1 VP02_BPPRD	P27378 bacterioph
16	40	45.5	245	1 UNG_MYCGE	P47343 mycoplasma
17	40	45.5	286	1 CAZ1_CHICK	P13127 gallus gall
18	40	45.5	455	1 SYPC_METJA	Q59635 methanococc
19	40	45.5	474	1 DNB2_ADE41	P11807 human adeno
20	40	45.5	484	1 DNB2_ADE12	P04498 human adeno
21	40	45.5	776	1 AD07_MACFA	Q28475 macaca fasc
22	40	45.5	1211	1 ATH1_YEAST	P48016 saccharomyc
23	39.5	44.9	1391	1 LYS2_CANAL	Q12572 candida alb
24	39	44.3	78	1 R27A_ASPOF	P31753 asparagus o
25	39	44.3	79	1 R27A_HORVU	P22277 hordeum vul
26	39	44.3	79	1 R27A_MAIZE	P27923 zea mays (m
27	39	44.3	80	1 R27A_LYCES	P27083 lycopersico
28	39	44.3	81	1 27AB_ARATH	P59232 arabidopsis
29	39	44.3	81	1 27AC_ARATH	P59233 arabidopsis
30	39	44.3	212	1 BCAS3_MOUSE	Q9JJR5 mus musculu
31	39	44.3	420	1 Y461_SYNY3	Q55167 synechocyst
32	39	44.3	754	1 AD07_HUMAN	Q9H2U9 homo sapien
33	39	44.3	832	1 SM4B_HUMAN	Q9NPR2 homo sapien

34	39	44.3	1077	1 XINY_CLOTM	P51584 clostridium
35	39	44.3	1215	1 C1KA_BACTM	Q45715 bacillus th
36	38	43.2	322	1 PSA_BRARE	Q9YHV4 brachydanio
37	38	43.2	329	1 HEM2_SCHPO	P78974 schizosacch
38	38	43.2	365	1 GLN2_DROME	P20478 drosophila
39	38	43.2	367	1 GLNA_CABEL	P34497 caenorhabdi
40	38	43.2	373	1 GLNA_MOUSE	P15105 mus musculu
41	38	43.2	373	1 GLNA_RAT	P09606 rattus norv
42	38	43.2	399	1 GLN1_DROME	P20477 drosophila
43	38	43.2	401	1 COAT_PAV	Q9J7Z0 pariacoto v
44	38	43.2	403	1 GLNA_SQUAC	P41320 squaius aca
45	38	43.2	486	1 DNAB_RICPR	Q9ZD08 rickettsia

ALIGNMENTS

RESULT 1
FGL2_MOUSE
ID FGL2_MOUSE STANDARD; PRT; 432 AA.
AC P12804;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibroleukin precursor (Fibrinogen-like protein 2) (Prothrombinase)
DE (Cytotoxic T-lymphocyte specific protein).
GN FGL2 OR FIBLP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cytotoxic T-cell;
RX MEDLINE=87175527; PubMed=3550794;
RA Koyama T., Hall L.R., Haser W.G., Tonogawa S., Saito H.;
RT "Structure of a cytotoxic T-lymphocyte-specific gene shows a strong
RT homology to fibrinogen beta and gamma chains.";
RJ Proc. Natl. Acad. Sci. U.S.A. 84:1609-1613(1987).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=BALE/CJ; TISSUE=Peritoneal macrophage;
RX MEDLINE=95333285; PubMed=7609073;
RA Part R.L., Fung L., Reneker C., Myers-Nason N., Leibowitz J.L.,
RA Levy G.;
RT "Association of mouse fibrinogen-like protein with murine hepatitis
RT virus-induced prothrombinase activity.";
RJ J. Virol. 69:5033-5038(1995).
CC -!- FUNCTION: CONVERTS PROTHROMBIN TO THROMBIN.
CC -!- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC
CC T-CELLS.
CC -!- INDUCTION: IN MACROPHAGES, DURING INFECTION BY MOUSE HEPATITIS
CC VIRUS STRAIN 3 (MHV-3).
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16238; AAA37624.1; -;
CC EMBL; M15761; AAA37624.1; JOINED.
CC EMBL; S78773; AAB34823.1; -;
CC PIR; A27447; A27447.
CC HSSP; P02671; 1FZD.
CC MGD; MGI:103266; Fgl2.
CC InterPro; IPR002181; Fibrinogen_C.
CC Pfam; PF00147; fibrinogen_C; 1.

DR SMART; SMCC186; FBG; 1.
DR PROSITE; PS00514; FIBRIN AG_C_DOMAIN; 1.
KW T-cell; Cytolysis; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 FIBROLEUKIN.
FT DOMAIN 203 428 FIBRINOGEN C-TERMINAL.
FT DISULFID 206 235 BY SIMILARITY.
FT DISULFID 364 377 BY SIMILARITY.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 332 332 A -> G (IN REF. 2).
SQ SEQUENCE 432 AA; 48951 MW; 2B297F69CB4A782 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 432;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
:|||||||
Cb 357 DRYPSGNCGLYSSG 371

RESULT 2
FGL2 HUMAN STANDARD; PRT; 439 AA.
AC Q14314;
DT 16-OCT-2001 (Rel. 40, Created;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinoleukin precursor (Fibrinogen-like protein 2) (p749).
GN FGL2.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=95369700; PubMed=7642106;
RA Ruegg C., Pytela R.;
RT "Sequence of a human transcript expressed in T-lymphocytes and
RT encoding a fibrinogen-like protein."
RJ Gene 160:257-262(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Yuvaraj S., Liu M., Marsden P., Levy G.;
RT "Cloning and characterization of Hfgl2: the human counterpart to the
RT mouse gene Fgl2."
RJ Submitted (NOV-1998) to the EXEL/GenBank/DBC databases.
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLU-53.
RA Rieder M.J., Armet T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA Poel C.L., Ozuna M., Yi Q., Nickerson D.A.;
RJ Submitted (JAN-2002) to the EXBL/GenBank/DBC databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.J., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin I.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulianhy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=98309432; PubMed=9647217;
RA Marazzi S., Blum S., Hartmann R., Gundersen D., Schreyer M.,
RA Argaves S., von Fliedner V., Pytela R., Ruegg C.;
RT "Characterization of human fibrobleukin, a fibrinogen-like protein
RT secreted by T lymphocytes";
RJ J. Immunol. 161:138-147(1998).
CC -!- FUNCTION: MAY PLAY A ROLE IN PHYSIOLOGIC LYMPHOCYTE FUNCTIONS AT
CC MUCOSAL SITES.
CC -!- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC
CC T-CELLS.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EXBL; Z36531; CAA85298.1; -.
CC EXBL; AF1C4015; AAD10825.1; -.
CC EMBL; AF1C4014; AAD10825.1; JOINED.
CC EMBL; AF468959; AAL68855.1; -.
CC EMBL; BC033820; AAH33820.1; -.
CC PIR; I37391; I37391.
CC HSSP; P02671; 1FZD.
CC Genew; HGNC:3696; FGL2.
CC MIM; 605351; -.
CC GO; GO:0005576; C:extracellular; TAS.
CC GO; GO:0005973; C:fibrinogen beta chain; TAS.
CC GO; GO:0005974; C:fibrinogen gamma chain; TAS.
CC InterPro; IPR002181; Fibrinogen_C.
CC Pfam; PF00147; fibrinogen_C; 1.
CC SMART; SM00186; FBG; 1.
CC PROSITE; PS00514; FIBRIN AG_C_DOMAIN; 1.
KW T-cell; Glycoprotein; Signal; Polymorphism.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 439 FIBROLEUKIN.
FT DOMAIN 210 435 FIBRINOGEN C-TERMINAL.
FT DISULFID 213 242 BY SIMILARITY.
FT DISULFID 371 384 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 53 53 G -> E.
FT
SQ SEQUENCE 439 AA; 50228 MW; DF34656288E49E68 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 439;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
:|||||||
Cb 364 DRYPSGNCGLYSSG 378

RESULT 3

```
RELN_CHICK
ID RELN_CHICK STANDARD; PRT; 3209 AA.
AC O93574;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Reelin (EC 3.4.21.-) (Fragment).
GN RELN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Bernier B., Goffinet A.M.;
RT "Comparative study of reelin in vertebrates.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: Extracellular matrix serine protease that plays a role
CC in layering of neurons in the cerebral cortex and cerebellum.
CC Regulates microtubule function in neurons and neuronal migration.
CC Affects migration of sympathetic preganglionic neurons in the
CC spinal cord, where it seems to act as a barrier to neuronal
CC migration. Enzymatic activity is important for the modulation of
CC cell adhesion. Binding to the extracellular domains of lipoprotein
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
CC Dab1 and modulation of Tau phosphorylation (By similarity).
CC !- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By
CC similarity).
CC !- SUBCELLULAR LOCATION: Secreted (By similarity).
CC !- DOMAIN: The basic C-terminal region is essential for secretion (By
CC similarity).
CC !- SIMILARITY: BELONGS TO THE REELIN FAMILY.
CC !- SIMILARITY: Contains 8 EGF-like domains.
CC !- SIMILARITY: Contains 15 BNR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF090441; AAC35559.1; -.
DR HSSP; P05106; 1JY2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002860; GH_BNR.
DR InterPro; IPR006210; EGF.
DR Pfam; PFC2012; BNR; 15.
DR Pfam; PFC0008; EGF; 4.
DR SMART; SMC0181; EGF; 5.
DR PROSITE; PS00022; EGF 1; 7.
DR PROSITE; PS01186; EGF 2; 7.
KW Hydrolase; Serine protease; Developmental protein; Matrix protein;
KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat.
FT NON_TER 1
FT DOMAIN 418 449 EGF-LIKE 1.
FT DOMAIN 777 808 EGF-LIKE 2.
FT DOMAIN 1157 1190 EGF-LIKE 3.
FT DOMAIN 1513 1544 EGF-LIKE 4.
FT DOMAIN 1877 1909 EGF-LIKE 5.
FT DOMAIN 2226 2257 EGF-LIKE 6.
FT DOMAIN 2601 2632 EGF-LIKE 7.
FT DOMAIN 2976 3008 EGF-LIKE 8.
FT REPEAT 340 351 BNR 1.
FT REPEAT 546 557 BNR 2.
FT REPEAT 699 710 BNR 3.
FT REPEAT 904 915 BNR 4.
FT REPEAT 1070 1081 BNR 5.
FT REPEAT 1283 1294 BNR 6.
FT REPEAT 1434 1445 BNR 7.
FT REPEAT 1632 1643 BNR 8.
```

```
FT REPEAT 1791 1802 BNR 9.
FT REPEAT 1998 2009 BNR 10.
FT REPEAT 2147 2158 BNR 11.
FT REPEAT 2346 2357 BNR 12.
FT REPEAT 2526 2537 BNR 13.
FT REPEAT 2727 2738 BNR 14.
FT REPEAT 3111 3122 BNR 15.
FT DOMAIN ARG-RICH (BASIC).
FT CARBOHYD 5 5
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1195 1195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1348 1348 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1669 1669 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1893 1893 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2017 2017 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2065 2065 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2317 2317 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2710 2710 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2764 2764 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2821 2821 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2933 2933 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3160 3160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3187 3187 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 3209 AA; 361288 MW; 81A7B6676BCAA3D1 CRC64;

Query Match 52.3%; Score 46; DB 1; Length 3209;
Best Local Similarity 71.4%; Pred. No. 34;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNCGLY--YSS 14
DB 1722 YPGGNIGLYCPYSS 1735

RESULT 4
RELN_HUMAN
ID RELN_HUMAN STANDARD; PRT; 3460 AA.
AC P78539; Q9UDQ2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Reelin precursor (EC 3.4.21.-).
GN RELN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97202106; PubMed=9049633;
RA Desilva U., D'Arcangelo G., Braden V.V., Chen J., Miao G.G.,
RA Curran T., Green E.D.;
RT "The human reelin gene: isolation, sequencing, and mapping on
RT chromosome 7.";
RL Genome Res. 7:157-164(1997).
RN [2]
RP SEQUENCE OF 194-2556 FROM N.A.
RA Lamar B., Wamsley P., Gibson A., Maas J., Bauer C., Sapetti L.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP ALTERNATIVE SPLICING.
RX MEDLINE=99263436; PubMed=10328932;
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergeyck V.,
RA Goffinet A.M.;
RT "Evolutionarily conserved, alternative splicing of reelin during brain
RT development.";
RL Exp. Neurol. 156:229-238(1999).
RN [4]
RP DISEASE, AND TISSUE SPECIFICITY.
RX MEDLINE=99080080; PubMed=9861036;
```

RA Impagnatiello F., Guidotti A.R., Pesold C., Dwivedi Y., Caruncho H.,
RA Pisu M.G., Uzuncu D.P., Smalheiser N.R., Davis J.M., Pandey G.N.,
RA Pappas G.D., Tueting P., Sharma R.P., Costa E.,
RT "A decrease of reelin expression as a putative vulnerability factor in
RT schizophrenia.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:15718-15723(1998).
RN [5]
RN DISEASE.
RX MEDLINE=20428190; PubMed=10973257;
RA Hong S.E., Shugart Y.Y., Huang D.T., Shahwan S.A., Grant P.E.,
RA Hourihane J.O.B., Martin N.D.T., Walsh C.A.,
RT "Autosomal recessive lissencephaly with cerebellar hypoplasia is
RT associated with human REELN mutations.";
RL Nat. Genet. 26:93-96(2000).
RN [6]
RN ERRATUM.
RA Hong S.E., Shugart Y.Y., Huang D.T., Shahwan S.A., Grant P.E.,
RA Hourihane J.O.B., Martin N.D.T., Walsh C.A.,
RL Nat. Genet. 27:225-225(2001).
RN [7]
RN DISEASE.
RX MEDLINE=21217116; PubMed=11317216;
RA Persico A.M., D'Agruma L., Maiorano N., Totaro A., Militeri R.,
RA Bravaccio C., Wassink T.H., Schneider C., Melmed R., Trillo S.,
RA Montecchi F., Palermo M., Pascucci T., Puglisi-Allegra S.,
RA Reichelt K.-L., Conciatori M., Marins R., Quattrocchi C.C., Baldi A.,
RA Zelante L., Gasparini P., Keller F.,
RT "Reelin gene alleles and haplotypes as a factor predisposing to
RT autistic disorder.";
RL Mol. Psych. 6:150-159(2001).
CC -!- FUNCTION: Extracellular matrix serine protease that plays a role
CC in layering of neurons in the cerebral cortex and cerebellum.
CC Regulates microtubule function in neurons and neuronal migration.
CC Affects migration of sympathetic preganglionic neurons in the
CC spinal cord, where it seems to act as a barrier to neuronal
CC migration. Enzymatic activity is important for the modulation of
CC cell adhesion. Binding to the extracellular domains of lipoprotein
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
CC Dab1 and modulation of Tau phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P78509-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P78509-2; Sequence=VSP_005575;
CC Name=3;
CC IsoId=P78509-3; Sequence=VSP_005576;
CC -!- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis
CC by the Cajal-Retzius cells and other pioneer neurons located in
CC the telencephalic marginal zone and by granule cells of the
CC external granular layer of the cerebellum. In adult brain,
CC preferentially expressed in GABAergic interneurons of prefrontal
CC cortices, temporal cortex, hippocampus and glutamatergic granule
CC cells of cerebellum. Also expressed in fetal and adult liver.
CC -!- DEVELOPMENTAL STAGE: Expressed in fetal and postnatal brain and
CC liver. Expression in postnatal human brain is high in the
CC cerebellum.
CC -!- DOMAIN: The basic C-terminal region is essential for secretion (By
CC similarity).
CC -!- DISEASE: Defects in REELN are the cause of autosomal recessive
CC lissencephaly with cerebellar hypoplasia (Also known as Norman-
CC Roberts syndrome). Some patients also displayed persistent
CC lymphedema neonatally, and one showed accumulation of chylous
CC (i.e., fatty) ascites fluid.
CC -!- DISEASE: Defects in REELN may contribute to susceptibility to
CC schizophrenia. Expression of the protein is reduced (about 50%) in
CC patients with schizophrenia.
CC -!- DISEASE: Defects in REELN may predispose to autistic disorder. A
CC polymorphic GGC triplet repeat located in the 5'UTR region of REELN
CC (8 to 10 repeats) in the normal population is significantly

CC increased in autistic patients (4 to 23 additional repeats).
CC -!- SIMILARITY: BELONGS TO THE REELIN FAMILY.
CC -!- SIMILARITY: Contains 8 EGF-like domains.
CC -!- SIMILARITY: Contains 15 BNR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J79716; AAC51105.1; -;
DR EMBL; AC000121; AAB46357.1; -;
DR EMBL; AC006316; AAD29127.1; -;
DR HSSP; P05106; IJUV2.
DR Genew; HGNC:9957; REELN.
DR MIM; 600514; -;
DR MIM; 257320; -;
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002860; GH ENR.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR002861; Reeler.
DR Pfam; PF02012; BNR; 15.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF02014; Reeler; 1.
DR SMART; SM00181; EGF; 5.
DR PROSITE; PS00022; EGF 1; 7.
DR PROSITE; PS01186; EGF 2; 6.
ER Hydroxylase; Serine protease; Developmental protein; Matrix protein;
KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;
KW Alternative splicing; Lissencephaly.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 3460 REELIN.
FT DOMAIN 39 171 REELER.
FT DOMAIN 670 701 EGF-LIKE 1.
FT DOMAIN 1029 1060 EGF-LIKE 2.
FT DOMAIN 1408 1441 EGF-LIKE 3.
FT DOMAIN 1764 1795 EGF-LIKE 4.
FT DOMAIN 2128 2160 EGF-LIKE 5.
FT DOMAIN 2477 2508 EGF-LIKE 6.
FT DOMAIN 2852 2883 EGF-LIKE 7.
FT DOMAIN 3227 3259 EGF-LIKE 8.
FT REPEAT 592 603 BNR 1.
FT REPEAT 798 809 BNR 2.
FT REPEAT 951 962 BNR 3.
FT REPEAT 1156 1167 BNR 4.
FT REPEAT 1322 1333 BNR 5.
FT REPEAT 1534 1545 BNR 6.
FT REPEAT 1685 1696 BNR 7.
FT REPEAT 1883 1894 BNR 8.
FT REPEAT 2042 2053 BNR 9.
FT REPEAT 2249 2260 BNR 10.
FT REPEAT 2398 2409 BNR 11.
FT REPEAT 2597 2608 BNR 12.
FT REPEAT 2777 2788 BNR 13.
FT REPEAT 2978 2989 BNR 14.
FT REPEAT 3362 3373 BNR 15.
FT DOMAIN 3431 3460 ARG-RICH (BASIC).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1266 1266 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1599 1599 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1749 1749 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1920 1920 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2144 2144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2268 2268 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2316 2316 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2568 2568 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 2961 2961 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3015 3015 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3072 3072 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3184 3184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3411 3411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3438 3438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 3428 3429 Missing (in isoform 2).
FT VARSPLIC 3428 3460 /FTid=VSP_005575.
FT CONFLICT 752 752 Missing (in isoform 3).
FT SEQUENCE 3460 AA; 388399 MW; 3B2C0BA415AB93C1 CRC64;
SQ
Query Match 52.3%; Score 46; DB 1; Length 3460;
Best Local Similarity 71.4%; Pred. No. 37;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
CY 3 YPSGNCGLY--YSS 14
Db 1973 YPGCNIGLYCPYSS 1986
||| ||| |||
||| ||| |||
```

RESULT 5

```
RELN_MOUSE
ID RELN_MOUSE STANDARD; PRT; 3461 AA.
AC Q60841; Q9CUA6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Reelin precursor (EC 3.4.21.-) (Reeler protein).
GN RELN OR RL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RX MEDLINE=95231649; PubMed=7715726;
RA D'Arcangelo G., Miao G.G., Chen S.-C., Soares H.D., Morgan J.L.,
RA Curran T.;
RA "A protein related to extracellular matrix proteins deleted in the
RA mouse mutant reeler.";
RA Nature 374:719-723(1995).
RN [2]
SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=98086481; PubMed=9417911;
RA Royaux I., Lambert de Rouvroit C., D'Arcangelo G., Demirov D.,
RA Goffinet A.M.;
RA "Genomic organization of the mouse reelin gene.";
RA Genomics 46:240-250(1997).
RN [3]
SEQUENCE OF 2152-3461 FROM N.A. (ISOFORM 1).
RC STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=95375789; PubMed=7647795;
RA Hirotsune S., Takahara T., Sasaki N., Hirose K., Yoshiki A.,
RA Chashi T., Kusakabe M., Murakami Y., Muramatsu M., Watanabe S.,
RA Nakao K., Katsuki M., Hayashizaki Y.;
RA "The reeler gene encodes a protein with an EGF-like motif expressed by
RA pioneer neurons.";
RA Nat. Genet. 10:77-83(1995).
RN [4]
SEQUENCE OF 1044-3461 FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
```

```
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake C., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima C., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RA Nature 409:685-690(2001).
RN [5]
CHARACTERIZATION.
RX MEDLINE=97141547; PubMed=8987733;
RA D'Arcangelo G., Nakajima K., Miyata T., Ogawa M., Mikoshiba K.,
RA Curran T.;
RA "Reelin is a secreted glycoprotein recognized by the CR-50 monoclonal
RA antibody.";
RA J. Neurosci. 17:23-31(1997).
RN [6]
CHARACTERIZATION.
RX MEDLINE=21634904; PubMed=11689558;
RA Quattrocchi C.C., Wannenes F., Persico A.M., Ciafre S.A.,
RA D'Arcangelo G., Farace M.G., Keller F.;
RA "Reelin is a serine protease of the extracellular matrix.";
RA J. Biol. Chem. 277:303-309(2002).
RN [7]
TISSUE SPECIFICITY.
RX MEDLINE=97325946; PubMed=9182958;
RA Schiffmann S.N., Bernier B., Goffinet A.M.;
RA "Reelin mRNA expression during mouse brain development.";
RA Eur. J. Neurosci. 9:1055-1071(1997).
RN [8]
ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=99263436; PubMed=10328932;
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergcyck V.,
RA Goffinet A.M.;
RA "Evolutionarily conserved, alternative splicing of reelin during brain
RA development.";
RA Exp. Neurol. 156:229-238(1999).
RN [9]
BINDING TO VLDLR AND APOER2.
RX MEDLINE=20036019; PubMed=10571241;
RA Hiesberger T., Trommsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,
RA Cooper J.A., Herz J.;
RA "Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces
RA tyrosine phosphorylation of disabled-1 and modulates tau
RA phosphorylation.";
RA Neuron 24:481-489(1999).
RN [10]
FUNCTION.
RX MEDLINE=20359755; PubMed=10860573;
RA Yip J.W., Yip Y.P.L., Nakajima K., Capriotti C.;
RA "Reelin controls position of autonomic neurons in the spinal cord.";
RA Proc. Natl. Acad. Sci. U.S.A. 97:8612-8616(2000).
CC -!- FUNCTION: Extracellular matrix serine protease that plays a role
CC in layering of neurons in the cerebral cortex and cerebellum.
CC Regulates microtubule function in neurons and neuronal migration.
CC Affects migration of sympathetic preganglionic neurons in the
CC spinal cord, where it seems to act as a barrier to neuronal
CC migration. Enzymatic activity is important for the modulation of
CC cell adhesion. Binding to the extracellular domains of lipoprotein
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
CC Dab1 and modulation of Tau phosphorylation.
CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q60841-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q60841-2; Sequence=VSP_005577;
CC
```

CC Name=3;
CC IsoId=Q60841-3; Sequence=VSP_005578;
CC TISSUE SPECIFICITY: The major isoform 1 is neuron-specific. It is
CC abundantly produced during brain ontogenesis by the Cajal-Retzius
CC cells and other pioneer neurons located in the telencephalic
CC marginal zone and by granule cells of the external granular layer
CC of the cerebellum. Expression is located in deeper layers in the
CC developing hippocampus and olfactory bulb, low levels of
CC expression are also detected in the immature striatum. At early
CC developmental stages, expressed also in hypothalamic
CC differentiation fields, tectum and spinal cord. A moderate to low
CC level of expression occurs in the septal area, striatal fields,
CC habenular nuclei, some thalamic nuclei, particularly the lateral
CC geniculate, the retina and some nuclei of the reticular formation
CC in the central field of the medulla. Very low levels found in
CC liver and kidney. No expression in radial glial cells, cortical
CC plate, Purkinje cells and inferior olivary neurons. The minor
CC isoform 2 is only expressed in non-neuronal cells. The minor
CC isoform 3 is found in the same cells as isoform 1, but is almost
CC undetectable in retina and brain stem.
CC DEVELOPMENTAL STAGE: First detected at embryonic day 11.5.
CC Expression increases up to birth and remains high from post-natal
CC day 2 to 11 in both cerebellum and fore/midbrain. Expression
CC declines thereafter and is largely brain specific in the adult.
CC DOMAIN: The basic C-terminal region is essential for secretion.
CC PTM: N-glycosylated and to a lesser extent also O-glycosylated.
CC DISEASE: Defects in reelin are the cause of the autosomal recessive
CC reeler (rl) phenotype which is characterized by impaired motor
CC coordination, tremors and ataxia. Neurons in affected mice fail to
CC reach their correct locations in the developing brain, disrupting
CC the organization of the cerebellar and cerebral cortices and other
CC laminated regions.
CC SIMILARITY: BELONGS TO THE REELIN FAMILY.
CC SIMILARITY: Contains 8 EGF-like domains.
CC SIMILARITY: Contains 15 BNR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24703; AAB91599.1; -
CC EMBL; D63520; BAA09788.1; ALT INIT.
CC EMBL; AK017094; BAB30592.1; -
CC MGD; MG1:103022; Reelin.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0016477; P:cell migration; IMP.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR002860; GH_BNR.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR002861; Reeler.
CC Pfam; PF02012; BNR; 15.
CC Pfam; PFC0008; EGF; 5.
CC Pfam; PF02014; Reeler; 1.
CC SMART; SM00181; EGF; 5.
CC PROSITE; PS00022; EGF_1; 7.
CC PROSITE; PS01186; EGF_2; 6.
KW Hydrolase; Serine protease; Developmental protein; Matrix protein;
KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;
KW Alternative splicing.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 3461 REELIN.
FT DOMAIN 40 172 REELER.
FT DOMAIN 671 702 EGF-LIKE 1.
FT DOMAIN 1030 1061 EGF-LIKE 2.
FT DOMAIN 1409 1442 EGF-LIKE 3.
FT DOMAIN 1765 1796 EGF-LIKE 4.
FT DOMAIN 2129 2161 EGF-LIKE 5.
FT DOMAIN 2478 2509 EGF-LIKE 6.
FT DOMAIN 2853 2884 EGF-LIKE 7.

FT DOMAIN 3228 3260 EGF-LIKE 8.
FT REPEAT 593 604 BNR 1.
FT REPEAT 799 810 BNR 2.
FT REPEAT 952 963 BNR 3.
FT REPEAT 1157 1168 BNR 4.
Query Match 52.3%; Score 46; DB 1; Length 3461;
Best local Similarity 71.4%; Pred. NO. 37;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 3 YPSGNCGLY--YSS 14
DB 1974 YPGNGIGLYCPYSS 1987
RESULT 6
REELIN_RAT
ID REELIN_RAT STANDARD; PRT; 3462 AA.
AC PS8751;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Reelin precursor (EC 3.4.21.-).
GN REELN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RA Kikkawa S., Terashima T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBC databases.
RN [2]
RP ALTERNATIVE SPLICING.
RX MEDLINE=99263436; PubMed=10328932;
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergeyck V.,
RA Goffinet A.M.;
RT "Evolutionarily conserved, alternative splicing of reelin during brain
RT development."
RL Exp. Neurol. 156:229-238 (1999).
CC -!- FUNCTION: Extracellular matrix serine protease that plays a role
CC in layering of neurons in the cerebral cortex and cerebellum.
CC Regulates microtubule function in neurons and neuronal migration.
CC Affects migration of sympathetic preganglionic neurons in the
CC spinal cord, where it seems to act as a barrier to neuronal
CC migration. Enzymatic activity is important for the modulation of
CC cell adhesion. Binding to the extracellular domains of lipoprotein
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
CC Dab1 and modulation of Tau phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=PS8751-1; Sequence=Displayed;
CC Name=2;
CC IsoId=PS8751-2; Sequence=VSP_005579;
CC Name=3;
CC IsoId=PS8751-3; Sequence=VSP_005580;
CC TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis
CC by the Cajal-Retzius cells and other pioneer neurons located in
CC the telencephalic marginal zone and by granule cells of the
CC external granular layer of the cerebellum.
CC DOMAIN: The basic C-terminal region is essential for secretion (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE REELIN FAMILY.
CC -!- SIMILARITY: Contains 8 EGF-like domains.
CC -!- SIMILARITY: Contains 15 BNR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AB049473; BAB78470.1; -.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002860; GH_BNR.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002861; Reeler.
DR Pfam; PF02012; BNR; 15.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF02014; Reeler; 1.
DR SMART; SM00181; EGF; 6.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 6.
KW Hydrolase; Serine protease; Developmental protein; Matrix protein;
KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;
KW Alternative splicing.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 3462 REELIN.
FT DOMAIN 41 173 REELER.
FT DOMAIN 672 703 EGF-LIKE 1.
FT DOMAIN 1031 1062 EGF-LIKE 2.
FT DOMAIN 1410 1443 EGF-LIKE 3.
FT DOMAIN 1766 1797 EGF-LIKE 4.
FT DOMAIN 2130 2162 EGF-LIKE 5.
FT DOMAIN 2479 2510 EGF-LIKE 6.
FT DOMAIN 2854 2885 EGF-LIKE 7.
FT DOMAIN 3229 3261 EGF-LIKE 8.
FT REPEAT 594 605 BNR 1.
FT REPEAT 800 811 BNR 2.
FT REPEAT 953 964 BNR 3.
FT REPEAT 1158 1169 BNR 4.
FT REPEAT 1324 1335 BNR 5.
FT REPEAT 1536 1547 BNR 6.
FT REPEAT 1687 1698 BNR 7.
FT REPEAT 1885 1896 BNR 8.
FT REPEAT 2044 2055 BNR 9.
FT REPEAT 2251 2262 BNR 10.
FT REPEAT 2400 2411 BNR 11.
FT REPEAT 2599 2610 BNR 12.
FT REPEAT 2779 2790 BNR 13.
FT REPEAT 2980 2991 BNR 14.
FT REPEAT 3364 3375 BNR 15.
FT DOMAIN 3433 3462 ARG-RICH (BASIC).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1448 1448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1601 1601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1922 1922 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2146 2146 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2270 2270 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2318 2318 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2570 2570 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2963 2963 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3017 3017 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3074 3074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3186 3186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3413 3413 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3440 3440 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 3430 3431 Missing (in isoform 2).
FT VARSPLIC 3430 3462 /FTId=VSP_005579.
FT VARSPLIC 3430 3462 Missing (in isoform 3).
FT VARSPLIC 3430 3462 /FTId=VSP_005580.
FT SEQUENCE 3462 AA; 387525 MW; FCCF89BC90E035F6 CRC64;

Query Match 52.3%; Score 46; DB 1; Length 3462;
Best Local Similarity 71.4%; Pred.No. 37;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 3 YPSGNCGLY--YSS 14
DB 1975 YPGGNIGLYCPYSS 1988
| | | | | | | | | |
RESULT 7
BCA3_HUMAN STANDARD; PRT; 210 AA.
ID BCA3_HUMAN
AC Q9NO31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proline-rich protein BCA3 (Breast cancer associated gene 3 protein).
GN BCA3 OR C11ORF17.
OS Homo sapiens (Humar).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21418998; PubMed=11528127;
RA Amid C., Bahr A., Mujica A., Sampson N., Bikar S.E., Winterpacht A.,
RA Zabel B., Hankeln T., Schmidt E.R.;
RT "Comparative genomic sequencing reveals a strikingly similar
RT architecture of a conserved syntenic region on human chromosome
RT 11p15.3 (including gene ST5) and mouse chromosome 7.";
RJ Cytogenet. Cell Genet. 93:284-290(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Kitching R., Li H., Wong M.J., Kanaganayakam S., Beatty B., Kahn H.,
RA Seth A.;
RT "Characterization of a novel human breast cancer associated gene
RT (BCA3) encoding an alternatively spliced proline rich protein.";
RJ Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

Query Match 51.1%; Score 45; DB 1; Length 210;
Best Local Similarity 66.7%; Pred.No. 31;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 YPSGNCGLYSS 14
DB 95 YTSSQCGKYSS 106
| | | | | | | | | |
RESULT 8
CAZ3_HUMAN STANDARD; PRT; 299 AA.
ID CAZ3_HUMAN
AC Q96X2; Q969J0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE F-actin capping protein alpha-3 subunit (CapZ alpha-3) (Germ cell-
DE specific protein 3).
GN CAPZA3 OR CAPAA3 OR GSG3.
OS Homo sapiens (Humar).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers C.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs,"
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.N., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.C., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Yadan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: F-ACTIN CAPPING PROTEINS BIND IN A CA(2+)-INDEPENDENT
CC MANNER TO THE FAST GROWING ENDS OF ACTIN FILAMENTS (BARBED END).
CC THEREBY BLOCKING THE EXCHANGE OF SUBUNITS AT THESE ENDS. UNLIKE
CC OTHER CAPPING PROTEINS (SUCH AS GELSOLIN AND SEVERIN), THESE
CC PROTEINS DO NOT SEVER ACTIN FILAMENTS. MAY PLAY A ROLE IN THE
CC MORPHOGENESIS OF SPERMATID.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT (BY
CC SIMILARITY).
CC -!- TISSUE SPECIFICITY: Exclusively expressed in the testis.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN 24-DAY-OLD AND ADULT TESTIS, BUT
CC NOT IN 4-, 10- AND 16-DAY-OLD TESTIS.
CC -!- SIMILARITY: BELONGS TO THE F-ACTIN CAPPING PROTEIN ALPHA SUBUNIT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D87471; BAAL3433.1; -;
DR EMBL; AB026984; BAAB1887.1; -;
DR EMBL; AK016391; BAB30213.1; -;
DR EMBL; BCC49620; AAB49620.1; -;
DR MGD; MGI:106221; Cappa3.
DR InterPro; IPR002189; F-actin_cap_A.
DR Pfam; PFC1267; F-actin_cap_A; 1.
DR PRINTS; PR00191; FACTINCAPA.
DR PRODom; PD006960; F-actin_cap_A; 1.
DR PROSITE; PS00748; F_ACTIN_CAPPING_A_1; FALSE_NEG.
DR PROSITE; PS00749; F_ACTIN_CAPPING_A_2; 1.
KW Actin-binding; Capping protein; Multigene family.
FT CONFLICT 117 117 L -> I (IN REF. 3).
FT CONFLICT 225 225 F -> Y (IN REF. 3).
SQ SEQUENCE 299 AA; 34952 MW; 8C87579313F233C6 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 299;
Best Local Similarity 60.0%; Pred. No. 9.3;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSGNCGL 10
Db 131 DHPNGNCNV 140
|:|:|:|:
RESULT 10
CAZ3_RAT STANDARD; PRT; 299 AA.
ID CAZ3_RAT
AC Q3WJV6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE F-actin capping protein alpha-3 subunit (CapZ alpha-3).
GN CAPZA3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98069272; PubMed=9406198;
RA Hurst S., Howes E.A., Coadwell J., Jones R.;
RT "Expression of a testis-specific putative actin-capping protein
RT associated with the developing acrosome during rat spermiogenesis,"
RL Mol. Reprod. Dev. 49:81-91 (1998).
CC -!- FUNCTION: F-ACTIN CAPPING PROTEINS BIND IN A CA(2+)-INDEPENDENT
CC MANNER TO THE FAST GROWING ENDS OF ACTIN FILAMENTS (BARBED END).
CC THEREBY BLOCKING THE EXCHANGE OF SUBUNITS AT THESE ENDS. UNLIKE
CC OTHER CAPPING PROTEINS (SUCH AS GELSOLIN AND SEVERIN), THESE
CC PROTEINS DO NOT SEVER ACTIN FILAMENTS. MAY PLAY A ROLE IN THE
CC MORPHOGENESIS OF SPERMATID.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT (BY
CC SIMILARITY).
CC -!- TISSUE SPECIFICITY: Exclusively expressed in the testis.
CC -!- SIMILARITY: BELONGS TO THE F-ACTIN CAPPING PROTEIN ALPHA SUBUNIT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y12538; CAA73137.1; -;
DR InterPro; IPR002189; F-actin_cap_A.
DR Pfam; PFC1267; F-actin_cap_A; 1.
DR PRINTS; PR00191; FACTINCAPA.
DR PRODom; PD006960; F-actin_cap_A; 1.
DR PROSITE; PS00748; F_ACTIN_CAPPING_A_1; FALSE_NEG.
DR PROSITE; PS00749; F_ACTIN_CAPPING_A_2; 1.
KW Actin-binding; Capping protein; Multigene family.
SQ SEQUENCE 299 AA; 35007 MW; 3D7530885CF79B55 CRC64;
Query Match 48.9%; Score 43; DB 1; Length 299;
Best Local Similarity 60.0%; Pred. No. 9.3;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 DRYPSGNCGL 10
Db 131 DHPNGNCNV 140
|:|:|:|:
RESULT 11
CHA9_LYMDI STANDARD; PRT; 121 AA.
ID CHA9_LYMDI
AC P50603;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Chorion class A proteins LD9 (Fragment).

```
OS Lymantria dispar (Gypsy moth).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
CC Lymantriidae; Lymantria.
CX NCBI_TaxID=13123;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Choriogenic follicles;
RX MEDLINE=95018300; PubMed=7932786;
RA Lecierc R.F., Regier J.C.;
RT "Evolution of chorion gene families in lepidoptera: characterization
RT of 15 cDNAs from the gypsy moth.";
RL J. Mol. Evol. 39:244-254(1994).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
CC GYPSY MOTH.
CC -!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH
CC BELONG CLASSES A, CA AND HCA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04661; AAA67861.1; -
KW Eggshell; Chorion; Repeat; Multigene family.
FT NON TER 1
SQ SEQUENCE 121 AA; 11617 MW; 528E28501ED06617 CRC64;

Query Match 46.6%; Score 41; DB 1; Length 121;
Best Local Similarity 66.7%; Pred. No. 7.8;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RYPSGNCGL 10
DB 4: RYPAGACGI 49

RESULT 12
YK78_AERPE
ID YK78_AERPE STANDARD; PRT; 437 AA.
AC Q9YA60;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein APE2078.
GN APE2078.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382956;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Arkai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1";
RL DNA Res. 6:83-101(1999).
CC -!- SIMILARITY: Belongs to the ubid family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000063; BAA81089.1; -
DR PIR; A72513; A72513.
DR InterPro; IPR002830; carboxylase.
DR Pfam; PF01977; Ubid; 1.
DR TIGRFAFS; TIGR00148; TIGR00148; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 437 AA; 47635 MW; 5055140A80A2D602 CRC64;

Query Match 46.6%; Score 41; DB 1; Length 437;
Best Local Similarity 57.1%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 RYPSGNCGLYSSG 15
DB 112 RFYEGEAGLYLSSG :25

RESULT 13
CTA0_HUMAN
ID CTA0_HUMAN STANDARD; PRT; 488 AA.
AC Q96NM4; Q96IC9; Q9BQN5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein C20orf100.
GN C20orf100.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiya H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Ota T.,
RA Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K.,
RA Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y.,
RA Sekine M., Kikuchi H., Kanda K., Wagatsuma M., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones X., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhali P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leheslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.B., McConnachie L.J., McIlroy K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.F., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston C.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
```

```
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.B., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN {}
RP SEQUENCE FROM N.A. (ISOFORM 2);
RC TISSUE=Muscle;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek C.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M., Skalska U., Smalius D.E.,
RA Schnorch A., Schein J.E., Jones S.G.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96NM4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96NM4-2; Sequence=VSP_002187;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Contains 1 HMG box domain.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-52 is the initiator.
CC -!- CAUTION: Ref.1 (CAC36288) sequence differs from that shown due to
CC erroneous gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
QY 3 YPSGNCGL 10
Db 466 YPSGECGI 473
RESULT 14
DNB2_ADEC7 STANDARD; PRT; 517 AA.
AC P04497;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Early E2A DNA-binding protein.
GN DBP.
OS Human adenovirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10519;
RN {}
RP SEQUENCE FROM N.A.
RX MEDLINE=84185604; PubMed=6325415;
RA Quinn C.O., Kitchingman G.R.;
RT "Sequence of the DNA-binding protein gene of a human subgroup B
RT adenovirus (type 7). Comparisons with subgroup C (type 5) and
RT subgroup A (type 12).";
RL C. Biol. Chem. 259:5003-5009(1984).
CC -!- FUNCTION: BINDS COOPERATIVELY SINGLE-STRANDED DNA IN A SEQUENCE-
CC INDEPENDENT MANNER. INVOLVED IN DNA-REPLICATION, REGULATION OF
CC MRNA FORMATION, AND HOST-RANGE SPECIFICITY. ZINC IS REQUIRED
CC FOR DNA BINDING.
CC -!- SUBCELLULAR LOCATION: NUCLEAR. ACCUMULATES IN INFECTED CELLS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-9.
RX MEDLINE=91306449; PubMed=1853567;
RA Bamford J.K.H., Haenninen A.-L., Pakula T.M., Ojala P.M.,
RA Kalkinen N., Frilander M., Bamford D.H.;
RT "Genome organization of membrane-containing bacteriophage PRD1.";
RL Virology 183:658-676(1991).
CC -!- FUNCTION: THE LIFE CYCLE OF THE PHAGE BEGINS WHEN THE PHAGE
CC ADSORBS TO ITS RECEPTOR ON THE SURFACE OF THE HOST VIA THE
CC ADSORPTION PROTEIN P2 AND INJECTS ITS DNA INTO THE HOST CYTOPLASM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M69077; AAA32458.1; -;
DR PIR; D40477; WMBPQ2.
FT INIT MET C
SQ SEQUENCE 590 AA; 63690 MW; 5ACA024C92B8F6BC CRC64;

Query Match 45.6%; Score 41; DB 1; Length 590;
Best Local Similarity 46.7%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
| : ||| : |
Db 407 DEWVANNCGLFPMMSG 421

Search completed: November 5, 2003, 16:46:21
Job time : 2.66079 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 5, 2003, 16:37:23 ; Search time 1.99238 seconds
(without alignments)
1952.598 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL23:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	357	11 Q92PT7	Q92PT7 rattus norv
2	88	100.0	442	6 Q8YIP7	Q8YIP7 sus scrofa
3	56	63.6	220	5 Q8T8A2	Q8T8A2 ciona savig
4	56	63.6	652	5 Q9NDQ1	Q9NDQ1 ciona intes
5	46	52.3	1785	4 Q9Y211	Q9Y211 homo sapien
6	46	52.3	1785	4 Q9Y4V9	Q9Y4V9 homo sapien
7	46	52.3	2403	4 Q9UGM2	Q9UGM2 homo sapien
8	46	52.3	2412	4 Q9UJ57	Q9UJ57 homo sapien
9	46	52.3	2413	4 Q96DU4	Q96DU4 homo sapien
10	46	52.3	2413	4 Q9UKJ4	Q9UKJ4 homo sapien
11	46	52.3	2426	4 Q9UGM3	Q9UGM3 homo sapien
12	45	51.1	183	4 Q8TAE0	Q8TAE0 homo sapien
13	45	51.1	210	4 Q8NBS2	Q8NBS2 homo sapien
14	45	51.1	210	4 Q8TAC6	Q8TAC6 homo sapien
15	44.5	50.6	457	11 Q8R0Z6	Q8R0Z6 mus musculu
16	44	50.0	217	7 Q9GJH7	Q9GJH7 salmo trutt

17	44	50.0	259	5 Q76468	Q76468 cryptospori
18	44	50.0	292	5 Q9VDM9	Q9VDM9 drosophila
19	44	50.0	297	10 Q8H358	Q8H358 oryza sativ
20	44	50.0	472	17 Q8TZD1	Q8TZD1 pyrococcus
21	44	50.0	544	5 Q8SW93	Q8SW93 encephalite
22	43	48.9	225	16 Q9PQ63	Q9PQ63 ureaplasma
23	43	48.9	299	11 Q9D4N3	Q9D4N3 mus musculu
24	43	48.9	339	5 Q21788	Q21788 caenorhabdi
25	43	48.9	456	16 Q9Z619	Q9Z619 cnlomydia p
26	43	48.9	1007	10 Q8VYR8	Q8VYR8 arabidopsis
27	43	48.9	1082	10 Q8WCUC	Q8WCUC sorghum bic
28	43	48.9	1244	5 Q8IDP9	Q8IDP9 plasmodium
29	43	48.9	2862	16 Q8R874	Q8R874 thermoaer
30	42.5	48.3	178	2 Q87568	Q87568 helicobacte
31	42.5	48.3	696	16 Q9Z390	Q9Z390 helicobacte
32	42	47.7	664	5 Q16977	Q16977 caenorhabdi
33	42	47.7	771	5 Q9NGD0	Q9NGD0 leishmania
34	42	47.7	771	5 Q8T7E0	Q8T7E0 leishmania
35	42	47.7	873	5 Q8WTL6	Q8WTL6 caenorhabdi
36	41.5	47.2	470	4 Q9BZZ0	Q9BZZ0 homo sapien
37	41.5	47.2	470	4 Q8N199	Q8N199 homo sapien
38	41	46.6	154	10 Q9FF72	Q9FF72 arabidopsis
39	41	46.6	191	10 Q8GJV9	Q8GJV9 populus alb
40	41	46.6	273	5 Q20000	Q20000 caenorhabdi
41	41	46.6	283	11 Q8BSX5	Q8BSX5 mus musculu
42	41	46.6	283	11 Q8BR11	Q8BR11 mus musculu
43	41	46.6	351	10 Q8W006	Q8W006 gelidium cr
44	41	46.6	368	16 Q9PMM4	Q9PMM4 campylobact
45	41	46.6	388	5 Q9U307	Q9U307 caenorhabdi

ALIGNMENTS

RESULT 1
Q9EPT7 ID Q9EPT7 PRELIMINARY; PRT; 357 AA.
AC Q9EPT7; STRAIN=Sprague-Dawley;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Prothrombinase FGL2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Rychlik D.F., Chien E., Philippe M.;
RT "FGL2 Expression in the Sprague-Dawley Rat."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF323608; AAG42269.1; -
DR HSSP; PC2671; IFZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF0C147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 357 AA; 40966 MW; 31183DD9A02EBBA9 CRC64;

Query Match 100.0%; Score 88; DB 11; Length 357;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
|||
DB 282 DRYPSGNCGLYSSG 296

RESULT 2
Q8MIP7 ID Q8MIP7 PRELIMINARY; PRT; 442 AA.
AC Q8MIP7;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Fibrinogen-like protein 2.
GN FGL2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Chanekar A., Liu H., Grant D.R., Levy G.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY112657; RAM52324.1; -.
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; Fibrinogen_C.1.
DR SMART: SM00186; FBG; 1.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 442 AA; 50579 MW; 67800D67AEDFF899 CRC64;

Query Match 100.0%; Score 88; DB 6; Length 442;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYYSSG 15
|||...|||
Db 367 DRYPSGNCGLYYSSG 381

RESULT 3
Q8T8A2 PRELIMINARY; PRT; 220 AA.
AC Q8T8A2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Fibrinogen-like protein (Fragment).
GN CS-FIBRINOGEN-LIKE.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920613; PubMed=11923208;
RA Imai K.S., Satoh N., Satou Y.;
RT "Early embryonic expression of FGF4/6/9 gene and its role in the
induction of mesenchyme and notochord in Ciona savignyi embryos.";
RL Development 129:1729-1738(2002).
DR EMBL: AB073373; BAB88674.1; -.
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; fibrinogen_C.1.
DR SMART: SM00186; FBG; 1.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.
FT NON TER 1
SQ SEQUENCE 220 AA; 25758 MW; 4160FAE727F3ED06 CRC64;

Query Match 63.6%; Score 56; DB 5; Length 220;
Best Local Similarity 66.7%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 5; Indels 3; Gaps 0;

QY 1 DRYPSGNCGLYYSSG 15
|||...|||
Cb 143 DGYALGNCGRYYRSG 157

RESULT 4
Q9NDQ1 PRELIMINARY; PRT; 652 AA.
AC Q9NDQ1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

Fibrinogen-like protein.
CI-FIBRN.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
RA Satoh N.;
RT "Characterization of Brachyury downstream notochord genes in the Ciona
intestinalis embryo.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB036849; BAB00626.1; -.
DR HSSP: P02671; PF2D.
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; fibrinogen_C.1.
DR SMART: SM00186; FBG; 1.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 652 AA; 73252 MW; A492BA325162F0E0 CRC64;

Query Match 63.6%; Score 56; DB 5; Length 652;
Best Local Similarity 66.7%; Pred. No. 0.73;
Matches 10; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

QY 1 DRYPSGNCGLYYSSG 15
|||...|||
Cb 574 DGYALGNCGRYYRSG 588

RESULT 5
Q9Y211 PRELIMINARY; PRT; 1785 AA.
AC Q9Y211;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DMBT1.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20017478; PubMed=10551316;
RA Takeshita H., Sato M., Shiwaku H.O., Semba S., Sakurada A., Hoshi M.,
RA Hayashi Y., Tagawa Y., Ayabe H., Horii A.;
RT "Expression of the DMBT1 gene is frequently suppressed in human lung
cancer.";
RL Jpn. J. Cancer Res. 90:903-908(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL: AB020851; BAA78577.1; -.
DR EMBL: AB020812; BAA78577.1; JOINED.
DR EMBL: AB020813; BAA78577.1; JOINED.
DR EMBL: AB020814; BAA78577.1; JOINED.
DR EMBL: AB020815; BAA78577.1; JOINED.
DR EMBL: AB020816; BAA78577.1; JOINED.
DR EMBL: AB020817; BAA78577.1; JOINED.
DR EMBL: AB020818; BAA78577.1; JOINED.
DR EMBL: AB020819; BAA78577.1; JOINED.
DR EMBL: AB020820; BAA78577.1; JOINED.
DR EMBL: AB020821; BAA78577.1; JOINED.
DR EMBL: AB020822; BAA78577.1; JOINED.
DR EMBL: AB020823; BAA78577.1; JOINED.
DR EMBL: AB020824; BAA78577.1; JOINED.
DR EMBL: AB020825; BAA78577.1; JOINED.
DR EMBL: AB020826; BAA78577.1; JOINED.
DR EMBL: AB020827; BAA78577.1; JOINED.
DR EMBL: AB020828; BAA78577.1; JOINED.
DR EMBL: AB020829; BAA78577.1; JOINED.
DR EMBL: AB020830; BAA78577.1; JOINED.
DR EMBL: AB020831; BAA78577.1; JOINED.

DR EMBL; AB020832; BAA78577.1; JOINED.
DR EMBL; AB020833; BAA78577.1; JOINED.
DR EMBL; AB020834; BAA78577.1; JOINED.
DR EMBL; AB020835; BAA78577.1; JOINED.
DR EMBL; AB020836; BAA78577.1; JOINED.
DR EMBL; AB020837; BAA78577.1; JOINED.
DR EMBL; AB020838; BAA78577.1; JOINED.
DR EMBL; AB020839; BAA78577.1; JOINED.
DR EMBL; AB020840; BAA78577.1; JOINED.
DR EMBL; AB020841; BAA78577.1; JOINED.
DR EMBL; AB020842; BAA78577.1; JOINED.
DR EMBL; AB020843; BAA78577.1; JOINED.
DR EMBL; AB020844; BAA78577.1; JOINED.
DR EMBL; AB020845; BAA78577.1; JOINED.
DR EMBL; AB020846; BAA78577.1; JOINED.
DR EMBL; AB020847; BAA78577.1; JOINED.
DR EMBL; AB020848; BAA78577.1; JOINED.
DR EMBL; AB020849; BAA78577.1; JOINED.
DR EMBL; AB020850; BAA78577.1; JOINED.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 9.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 9.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 8.
DR PROSITE; PS00287; SRCR_2; 9.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
SQ SEQUENCE 1785 AA; 19391 MW; 38B2363F95226EB0 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 1785;
Best Local Similarity 64.3%; Pred. No. 93;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNCG--LYYSSG 15

Db 1134 PSSNCGGFLFYASG 1147

RESULT 6

Q9Y4V9
D Q9Y4V9 PRELIMINARY; PRT; 1785 AA.
AC Q9Y4V9;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE DMBT1/6kb.1 protein precursor.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=97434209; PubMed=9288095;
RA Mollenhauer J., Wiemann S., Scheurien W., Korn B., Hayashi Y.,
RA Wilgerbus K.K., von Deimling A., Poustka A.;
RT "DMBT1, a new member of the SRCR superfamily on chromosome 10q25.3-
q26.1 is deleted in malignant brain tumors.";
RL Nat. Genet. 17:32-39(1997).
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AJ000342; CAA04019.1; -
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 9.

DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 9.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 8.
DR PROSITE; PS00287; SRCR_2; 9.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1785 DMBT1/6KB.1 PROTEIN.
SQ SEQUENCE 1785 AA; 19391 MW; 7B1F8D47E4A82092 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 1785;
Best Local Similarity 64.3%; Pred. No. 93;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNCG--LYYSSG 15

Db 1134 PSSNCGGFLFYASG 1147

RESULT 7

Q9UGM2
ID Q9UGM2 PRELIMINARY; PRT; 2403 AA.
AC Q9UGM2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE DMBT1/8kb.1 protein.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=2065089; PubMed=10597221;
RA Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herbertz S.,
RA Madsen J., Kioschis P., Coy J.F., Poustka A.;
RT "The genomic structure of the DMBT1 gene: evidence for a region with
susceptibility to genomic instability.";
RL Oncogene 18:6233-6240(1999).
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AJ243224; CAB63942.1; -
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR TIGRFAMs; TIGR01443; intein_Cterm; 13.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 13.
DR PROSITE; PS00287; SRCR_2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
FT CHAIN 26 2403 DMBT1/8KB.1 PROTEIN.
SQ SEQUENCE 2403 AA; 259573 MW; 902622DE7399AD38 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 2403;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNCG--LYYSSG 15

Db 1752 PSSNCGGFLFYASG 1765

```
RESULT 8
Q9UJ57
ID Q9UJ57 PRELIMINARY; PRT; 2412 AA.
AC Q9UJ57;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DMBT1/8kb.2 protein precursor.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Trachea;
RX MEDLINE=99415938; PubMed=10485905;
RA Holmskov J., Mollenhauer J., Madsen J., Vitved L., Gronlund J.,
RA Tornoe I., Klien A., Reid K.B., Poustka A., Skjodt K.;
RT "Cloning of gp-340, a putative opsonin receptor for lung surfactant
protein D.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AJ243212; CAB56155.1; -.
DR HSSP; P29392; 1SFP.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR TIGRFAMs; TIGR01443; intein_Cterm; 13.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 13.
DR PROSITE; PS0287; SRCR_2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 2413 DMBT1/8KB.2 PROTEIN.
SQ SEQUENCE 2412 AA; 260568 MW; 3F630CCBFFB18EDC CRC64;

Query Match 52.3%; Score 46; DB 4; Length 2412;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNCG--LYYSSG 15
Db :762 PSSNCGGFLFYASG 1775

RESULT 9
Q96DU4
ID Q96DU4 PRELIMINARY; PRT; 2413 AA.
AC Q96DU4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DMBT1/8kb.2 protein precursor.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=99415938; PubMed=10485905;
RA Holmskov J., Mollenhauer J., Madsen J., Vitved L., Gronlund J.,
RA Tornoe I., Klien A., Reid K.B., Poustka A., Skjodt K.;
RT "Cloning of gp-340, a putative opsonin receptor for lung surfactant
protein D.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AF159456; AAD49696.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR TIGRFAMs; TIGR01443; intein_Cterm; 12.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 13.
```

```
RA Mollenhauer J.;
RT "Major subforms of DMBT1 are gastrointestinal mucins that display
extensive alternative splicing and differential protein targeting.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AJ297935; CAC44122.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR TIGRFAMs; TIGR01443; intein_Cterm; 13.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 13.
DR PROSITE; PS0287; SRCR_2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 2413 DMBT1/8KB.2 PROTEIN.
SQ SEQUENCE 2413 AA; 260600 MW; 99A449C5F4F60728 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 2413;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNCG--LYYSSG 15
Db :762 PSSNCGGFLFYASG 1775

RESULT 10
Q9UKZ4
ID Q9UKZ4 PRELIMINARY; PRT; 2413 AA.
AC Q9UKZ4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Gp-340 variant protein.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99415938; PubMed=10485905;
RA Holmskov J., Mollenhauer J., Madsen J., Vitved L., Gronlund J.,
RA Tornoe I., Klien A., Reid K.B., Poustka A., Skjodt K.;
RT "Cloning of gp-340, a putative opsonin receptor for lung surfactant
protein D.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AF159456; AAD49696.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR TIGRFAMs; TIGR01443; intein_Cterm; 12.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 13.
```

```
DR PROSITE; PS50287; SRCR 2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
SQ SEQUENCE 2413 AA; 260755 MW; 039544043CF463D4 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 2413;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYYSSG 15
DB 1762 PSSNCGGFLFYASG 1775

RESULT 11
Q9UGM3
ID Q9UGM3 PRELIMINARY; PRT; 2426 AA.
AC Q9UGM3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DI 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DMST1 prototype precursor.
GN DMST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20065089; PubMed=10597221;
RA Moellenhauer J., Holmskov U., Wiemann S., Krebs I., Herbertz S.,
RA Madsen J., Kischis P., Coy J.F., Poustka A.;
RT "The genomic structure of the DMST1 gene: evidence for a region with
RT susceptibility to genomic instability.";
RL Oncogene 18:6233-6240(1999).
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AJ243211; CAB63941.1; -.
DR Genew; HGNC:2926; DMST1.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF0530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 14.
DR SMART; SM00241; ZP; 1.
DR TIGRFAMs; TIGR01443; intein_cterm; 13.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; 13.
DR PROSITE; PS50287; SRCR_2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 2426 DMST1 PROTOTYPE.
SQ SEQUENCE 2426 AA; 262052 MW; 5A58FEC076FH7247 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 2426;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYYSSG 15
DB 1775 PSSNCGGFLFYASG 1788

RESULT 12
Q8TAE0
ID Q8TAE0 PRELIMINARY; PRT; 183 AA.
AC Q8TAE0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DI 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Keyt binding protein 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RA Li R., Han H., Wang J.;
RT "KBP, a novel protein interacting with LIM proteins KoyT.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF493784; AAM12863.1; -.
DR EMBL; AF493786; AAM12866.1; -.
SQ SEQUENCE 183 AA; 20207 MW; 872DD878B01C513B CRC64;

Query Match 51.1%; Score 45; DB 4; Length 183;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 3 YPSGNCGLYYSS 14
DB 95 YTSSQCGKYSS 106

RESULT 13
Q8NBS2
ID Q8NBS2 PRELIMINARY; PRT; 210 AA.
AC Q8NBS2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DI 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ90827.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Rattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDC human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075308; BAC11537.1; -.
KW Hypothetical protein.
SQ SEQUENCE 210 AA; 23084 MW; 0501CB50787367DC CRC64;

Query Match 51.1%; Score 45; DB 4; Length 210;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 3 YPSGNCGLYYSS 14
DB 95 YTSSQCGKYSS 106

RESULT 14
Q8TAC6
ID Q8TAC6 PRELIMINARY; PRT; 210 AA.
AC Q8TAC6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DI 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Keyt binding protein 1 (Chromosome 11 open reading frame 17).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

RC TISSUE=Lymph node;
RA Li R., Han H., Wang J.;
RT "KSP, a novel protein interacting with LIM protein Koyt.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF493783; AAM12862.1; -;
DR EMBL; AF493786; AAM12865.1; -;
DR EMBL; BC030996; AAH30996.1; -;
SQ SEQUENCE 210 AA; 23114 MW; FCCAAIF166C37967 CRC64;

Query Match 51.1%; Score 45; DB 4; Length 210;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 YPSGNCGLYSS 14
| : |||||
Db 95 YTSQCGKYSS 106

RESULT 15
Q8R0Z6
ID Q8R0Z6 PRELIMINARY; PRT; 457 AA.
AC Q8R0Z6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
ET 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
ET 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to angiopoietin-related protein 5 (ARP3).
GN 6330404ELLRIK CR ARP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Matsumoto S., Saito Y., Masuko Y., Yasunaga K., Oike Y., Suda T.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Oike Y., Suda T.;
RT "Molecular cloning of ARP3.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025904; AAH25904.1; -;
DR EMBL; AB054065; BAB91249.1; -;
DR MGD; MGI:1917976; 6330404ELLRIK.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C.1.
DR PROSITE; PS00514; FIBRINAG_C DOMAIN; 1.
SQ SEQUENCE 457 AA; 51095 MW; B7C4289E3FEC6C3E CRC64;

Query Match 50.6%; Score 44.5; DB 11; Length 457;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 DRYPSGNCGLYSSG 15
| : |||||
Db 391 DSY-SGNCALYHRGG 404

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 15:56:26 ; Search time 2.34053 Seconds
(without alignments)
609.684 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPSGNCGLYYSSG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 3.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	88	100.0	15	AAW88237	Human prothrombina
2	88	100.0	432	AAW88236	Mouse prothrombina
3	88	100.0	439	AAW88235	Human prothrombina
4	50	56.8	53	ABP34223	Human angiotensin
5	47	53.4	93	AAE03347	Human gene 1 encod
6	46	52.3	102	ABG25291	Kovel human diagno
7	46	52.3	666	AAW64590	Human SRCR protein
8	46	52.3	1785	AAW64591	Human SRCR protein
9	46	52.3	3460	ABB05007	Human reelin prote

10	46	52.3	3461	23	ABB05008	Mouse reelin prote
11	46	52.3	3461	23	ABB57065	Mouse ischaemic co
12	46	52.3	3470	22	ABG25297	Novel human diagno
13	45	51.1	123	22	AAU80079	Apoptin-associatin
14	45	51.1	126	22	AAU80078	Apoptin-associatin
15	45	51.1	126	23	AAO14806	Human apoptin-asso
16	45	51.1	158	21	AAE42752	Human ORF ORF2516
17	45	51.1	210	22	AAU80085	Apoptin-associatin
18	45	51.1	210	22	AAW93904	Human polypeptide,
19	45	51.1	210	23	ABP43815	CEGPI protein #1
20	45	51.1	210	24	AAE33640	Human BCA3 protein
21	44	50.0	292	22	ABB70950	Drosophila melanog
22	43	48.9	299	24	AAE32106	Human cytoskeleton
23	43	48.9	460	20	AAW34589	Chlamydia pneumonia
24	42.5	48.3	193	18	AAW55378	H. pylori ORF 07cp
25	42.5	48.3	193	23	AAW50313	Helicobacter pylor
26	42.5	48.3	493	18	AAW55517	H. pylori ORF 07ce
27	42.5	48.3	493	23	AAW50311	Helicobacter pylor
28	42.5	48.3	696	18	AAW55698	H. pylori ORF 14gp
29	42.5	48.3	696	20	AAW17207	H. pylori outer me
30	42.5	48.3	696	23	AAW50309	Helicobacter pylor
31	42.5	48.3	696	23	AAW50310	Helicobacter pylor
32	41.5	47.2	219	24	AAE32350	Human N18 fibrinog
33	41.5	47.2	279	22	AAW72625	Human angiotensin
34	41.5	47.2	342	22	AAW72626	Human angiotensin
35	41.5	47.2	470	20	AAW05398	Human TIE ligand N
36	41.5	47.2	470	21	AAW24389	Human PRO178 prote
37	41.5	47.2	470	22	AAW51330	Human NEW angiopo
38	41.5	47.2	470	22	AAW53066	Human angiogenesis
39	41.5	47.2	470	23	ABG80357	Human TIE ligand N
40	41.5	47.2	470	23	ABB95424	Human angiogenesis
41	41.5	47.2	470	23	AAE19827	Human TIE ligand N
42	41.5	47.2	470	23	ABB84818	Human PRO178 prote
43	41	46.6	39	17	AAW05340	Calisoga spider ve
44	41	46.6	39	17	AAW05341	Calisoga spider ve
45	41	46.6	80	17	AAW05343	Calisoga spider ve

ALIGNMENTS

RESULT 1
AAW88237
-D AAW88237 standard; Peptide; 15 AA.

XX	AAW88237;
AC	AAW88237;
XX	15-MAR-1999 (first entry);
DE	Human prothrombinase Fgl2 epitope.
XX	Prothrombinase; hfgl2; Fgl2; human; immune coagulation; antibody;
KW	inhibitor; infection; graft rejection; glomerulonephritis; cancer;
KW	gastrointestinal disease; foetal loss; therapy; vaccine; epitope.
XX	XX
OS	Homo sapiens.
XX	XX
PN	WC9851335-A1.
XX	XX
PD	19-NOV-1998.
XX	XX
PF	15-MAY-1998; 98WC-CA00475.
XX	XX
PR	10-OCT-1997; 97US-0061684.
FR	15-MAY-1997; 97US-0046537.
XX	XX
FA	(LEVY/) LEVY G.
XX	XX
PI	Levy G;
XX	XX
DR	WPI; 1999-059687/05.
XX	XX
PI	Modulating immune coagulation. - by using Fgl2 antibodies and

PT compounds, used to treat conditions including graft rejection and
PT foetal loss

FT foetal loss

PS
Claim: 4; Page 72; 105pp; English:

This peptide corresponds to amino acid residues 364-378 of human prothrombinase Fg12 (see AAW88235). A claimed method of preventing or treating a condition requiring a reduction in immune coagulation comprises administering an inhibitor of Fg12. The inhibitor is preferably an antibody that binds to the Fg12 epitope. The condition to be treated is graft rejection of foetal loss (claimed).

Sequence	15 AA:
50	

Query Match	100.0%;	Score 88;	DB 20;	Length 15;
Best Local Similarity	100.0%;	Pred. No. 1.2e-06;		
Matches 15; Conservative	0;	Mismatches 0;	Index's	

Qy 1 DRYPSGNCGLYSSG 15
| | | | | | | | | |
Db 1 DRYPSGNCGLYSSG 15

Db 1 DRYPSGNCGLYSSG 15

RESULT 2

AAW88236
ID AAW88236 standard: protein: 432 AA

AAW98236; AC

15-MAR-1999 (first entry)

DE Mouse prothrombinase Fc12 protein.

KW Prothrombinase; Fg12; mouse; immune coagulation; antibody;
KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;
KW gastrointestinal disease; foetal loss; therapy; vaccine.

SC
Yus SD.

Key	Location/Qualifiers
200	
201	
202	
203	
204	
205	
206	
207	
208	
209	
210	
211	
212	
213	
214	
215	
216	
217	
218	
219	
220	
221	
222	
223	
224	
225	
226	
227	
228	
229	
230	
231	
232	
233	
234	
235	
236	
237	
238	
239	
240	
241	
242	
243	
244	
245	
246	
247	
248	
249	
250	
251	
252	
253	
254	
255	
256	
257	
258	
259	
260	
261	
262	
263	
264	
265	
266	
267	
268	
269	
270	
271	
272	
273	
274	
275	
276	
277	
278	
279	
280	
281	
282	
283	
284	
285	
286	
287	
288	
289	
290	
291	
292	
293	
294	
295	
296	
297	
298	
299	
300	
301	
302	
303	
304	
305	
306	
307	
308	
309	
310	
311	
312	
313	
314	
315	
316	
317	
318	
319	
320	
321	
322	
323	
324	
325	
326	
327	
328	
329	
330	
331	
332	
333	
334	
335	
336	
337	
338	
339	
340	
341	
342	
343	
344	
345	
346	
347	
348	
349	
350	
351	
352	
353	
354	
355	
356	
357	
358	
359	
360	
361	
362	
363	
364	
365	
366	
367	
368	
369	
370	
371	
372	
373	
374	
375	
376	
377	
378	
379	
380	
381	
382	
383	
384	
385	
386	
387	
388	
389	
390	
391	
392	

Protein	Protein size	pI	pI ₂	pI ₃	pI ₄	pI ₅	pI ₆	pI ₇	pI ₈	pI ₉	pI ₁₀	pI ₁₁	pI ₁₂	pI ₁₃	pI ₁₄	pI ₁₅	pI ₁₆	pI ₁₇	pI ₁₈	pI ₁₉	pI ₂₀	pI ₂₁	pI ₂₂	pI ₂₃	pI ₂₄	pI ₂₅	pI ₂₆	pI ₂₇	pI ₂₈	pI ₂₉	pI ₃₀	pI ₃₁	pI ₃₂	pI ₃₃	pI ₃₄	pI ₃₅	pI ₃₆	pI ₃₇	pI ₃₈	pI ₃₉	pI ₄₀	pI ₄₁	pI ₄₂	pI ₄₃	pI ₄₄	pI ₄₅	pI ₄₆	pI ₄₇	pI ₄₈	pI ₄₉	pI ₅₀	pI ₅₁	pI ₅₂	pI ₅₃	pI ₅₄	pI ₅₅	pI ₅₆	pI ₅₇	pI ₅₈	pI ₅₉	pI ₆₀	pI ₆₁	pI ₆₂	pI ₆₃	pI ₆₄	pI ₆₅	pI ₆₆	pI ₆₇	pI ₆₈	pI ₆₉	pI ₇₀	pI ₇₁	pI ₇₂	pI ₇₃	pI ₇₄	pI ₇₅	pI ₇₆	pI ₇₇	pI ₇₈	pI ₇₉	pI ₈₀	pI ₈₁	pI ₈₂	pI ₈₃	pI ₈₄	pI ₈₅	pI ₈₆	pI ₈₇	pI ₈₈	pI ₈₉	pI ₉₀	pI ₉₁	pI ₉₂	pI ₉₃	pI ₉₄	pI ₉₅	pI ₉₆	pI ₉₇	pI ₉₈	pI ₉₉	pI ₁₀₀
Protein	Protein size	pI	pI ₂	pI ₃	pI ₄	pI ₅	pI ₆	pI ₇	pI ₈	pI ₉	pI ₁₀	pI ₁₁	pI ₁₂	pI ₁₃	pI ₁₄	pI ₁₅	pI ₁₆	pI ₁₇	pI ₁₈	pI ₁₉	pI ₂₀	pI ₂₁	pI ₂₂	pI ₂₃	pI ₂₄	pI ₂₅	pI ₂₆	pI ₂₇	pI ₂₈	pI ₂₉	pI ₃₀	pI ₃₁	pI ₃₂	pI ₃₃	pI ₃₄	pI ₃₅	pI ₃₆	pI ₃₇	pI ₃₈	pI ₃₉	pI ₄₀	pI ₄₁	pI ₄₂	pI ₄₃	pI ₄₄	pI ₄₅	pI ₄₆	pI ₄₇	pI ₄₈	pI ₄₉	pI ₅₀	pI ₅₁	pI ₅₂	pI ₅₃	pI ₅₄	pI ₅₅	pI ₅₆	pI ₅₇	pI ₅₈	pI ₅₉	pI ₆₀	pI ₆₁	pI ₆₂	pI ₆₃	pI ₆₄	pI ₆₅	pI ₆₆	pI ₆₇	pI ₆₈	pI ₆₉	pI ₇₀	pI ₇₁	pI ₇₂	pI ₇₃	pI ₇₄	pI ₇₅	pI ₇₆	pI ₇₇	pI ₇₈	pI ₇₉	pI ₈₀	pI ₈₁	pI ₈₂	pI ₈₃	pI ₈₄	pI ₈₅	pI ₈₆	pI ₈₇	pI ₈₈	pI ₈₉	pI ₉₀	pI ₉₁	pI ₉₂	pI ₉₃	pI ₉₄	pI ₉₅	pI ₉₆	pI ₉₇	pI ₉₈	pI ₉₉	pI ₁₀₀

Modified-size 228..231

```

F:
/note= "Asn is N-glycosylated"

```

```

FF Modified-site 256..259

```

```

E: /note= "Asn is N-glycosylated"
E:

```

FT	/note= "Asn is N-glycosylated"
FT	213..439
FT	Domain
FT	/note= "fibrinogen related domain"

WO9851335-A1

19-NCV-1998.

15-MAV-1998: 98WC-CA00475

)
 -
 .
 >
 >
 >
 ;
)

)
 :
)
 \

~
)
 \
 \
 #

 #
 #
 J
 -

)
 #

X
 # X

PR 15-MAY-1997:

PA (LEVY)

Jew G:

5
2
0
1
1
X
X

DR N-PSDB; AAV84140.

XX

PR Modulating immune coagulation - by using Fg12 antibodies and

PR compounds, used to treat conditions including graft rejection and

PR foetal loss

Claim 8: Page 70-71; :05pp; English.

XX Modulating immune coagulation - by using Fgl2 antibodies and
PT compounds, used to treat conditions including graft rejection and
PT foetal loss
PT
XX
PS Claim 5; Page 66-67; 105pp; English.
XX
CC This is the amino acid sequence of human prothrombinase Fgl2, as
CC predicted from Hgl2 DNA (see AAV84139). Fgl2 is a 70 kDa
CC transmembrane serine protease that has immune procoagulant activity.
CC The invention provides a method for inhibiting immune coagulation by
CC inhibiting the activity or expression of Fgl2. The method can be
CC used in vivo to treat a condition which requires a reduction in
CC immune coagulation such as bacterial and viral infections, cancer,
CC glomerulonephritis, a number of gastrointestinal diseases,
CC allograft and xenograft rejection and foetal loss. An Fgl2-specific
CC antibody, an Fgl2 antisense oligonucleotide, or a substance that
CC affects prothrombinase activity of a Fgl2 protein may be used to
CC treat a condition requiring a reduction in procoagulant activity.
CC A vaccine containing an Fgl2 protein or peptide is used for
CC prevention of graft rejection or foetal loss (claimed).

XX Sequence 439 AA;
SQ Query Match 100.0%; Score 88; DB 20; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

Oy 1 DRYPSGNCGLYSSG 15
|||:|||||
Db 364 DRYPSGNCGLYSSG 378

RESULT 4
ABP34223
ID ABP34223 standard; Protein; 53 AA.
XX
AC ABP34223;
XX
DT 08-JUL-2002 (first entry)
XX
DE Human angiotensin-like ORF3196 protein, SEQ ID NO:6392.
XX

KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulvar;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX Homo sapiens.
OS
XX WO200190366-A2.
PN
XX 29-NOV-2001.
PD
XX 24-MAY-2001; 2001WO-US17076.
PF

XX 24-MAY-2000; 2000US-206690P.
PR
XX (CURA-) CURAGEN CORP.
PA
XX Leach MD, Shimkets RA;
PI
XX WPI; 2002-106200/14.
XX
DR N-PSDB; ABN78249.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation -
XX
PS Claim 10; Page 1846; 2508pp; English.
XX
CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibitor activity, chemotactic/
CC chemokinetic activity, haemostatic activity, tumour inhibition activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX Sequence 53 AA;

Query Match 56.8%; Score 50; DB 23; Length 53;
Best Local Similarity 66.7%; Pred. No. 2.4;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 4 PSGNCGLYSSG 15
|||||:
Db 20 PSGNCALYHRGG 3;

RESULT 5
AAE03347
ID AAE03347 standard; peptide; 93 AA.

XX AAE03347;
AC
XX 10-AUG-2001 (first entry)
DI
XX Human gene 1 encoded secreted protein fragment, SEQ ID NO:122.

XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; uinary;

KW cell culture; chemotaxis; food additive; gene therapy;
KW binding partner identification.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 7 /label= Unknown
FT /note= "X equals stop translation"

XX W0200134800-A1.

XX 17-YAY-2001.

XX 08-NOV-2000; 2000WO-US30674.

XX 12-NOV-1999; 99US-0164750.

PR 30-JUN-2000; 2000US-0215128.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Ehner R, Fiscella M, Wei F;

XX WPI; 2001-329085/34.

XX New nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -

XX Disclosure; Page 501; 530pp; English.

XX AAD07705-AAD07759 represent cDNAs corresponding to 19 human secreted
CC protein genes, and AAE03292-AAE03346 represent the proteins they encode.
CC AAE03347-AAE03375 represent human secreted protein fragments or variants.
CC The genes and their secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 19 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumors, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein fragment referred to in the disclosure of the invention.

XX SQ Sequence 93 AA;

Query Match 53.4%; Score 47; DB 22; Length 93;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PSGNCGLYYSSG 15

Db |||||
32 PSGNCALYQRRG 43

RESULT 6

RESULT 7
AAW64590
ID AAW64590 standard; Protein; 666 AA.

ABG25291
ID ABG25291 standard; Protein; 102 AA.

XX AC ABG25291;

XX DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #25282.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WC250175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS89478.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX PS Claim 20; SEQ ID No 55650; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG0377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPC
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 102 AA;

Query Match 52.3%; Score 46; DB 22; Length 102;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNCGLY--YSS 14

Db |||||
86 YPGNIGLYCPYSS 99

XX AAW64590;
AC 23-OCT-1998 (first entry)
XX Human SRCR protein fragment.
XX Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
KW nervous system; medullo-blastoma; glioma; breast; detection;
KW autoantibody.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 3 /label= unknown
FT
XX WO9830687-A2.
PN 16-JUL-1998.
XX
PF 09-JAN-1998; 98WO-DE000096.
XX 18-JUL-1997; 97DE-1030997.
PR 09-JAN-1997; 97DE-1000519.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA Mollenhauer J, Poustka A;
XX WPI; 1998-399136/34.
DR N-PSDB; AAV49651.
XX Proteins containing scavenger receptor, cysteine rich domain -
PT useful for diagnosis and treatment of tumours
PT
XX Claim 1; Fig 1; 54pp; German.
XX This sequence represents a fragment of a human protein which contains a
CC SRCR (scavenger receptor, cysteine-rich) domain. The gene and encoded
CC protein can be used to diagnose or treat tumours, particularly of the
CC nervous system (medullo-blastoma or glioma) or breast. The DNA sequence
CC and probes derived from it, are used to identify genes that express
CC SRCR-domain containing proteins, to determine the form in which these
CC proteins exist and to assess the significance of individual forms on
CC cellular properties. The protein can be used to detect the presence of
CC autoantibodies, and Ab which regulate its expression.
XX
SQ Sequence 565 AA;
Query Match 52.3%; Score 46; DB 19; Length 666;
Best Local Similarity 64.3%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
OY 4 PSNCG--LYSSG 15
Db 183 PSSNCGGFLFYASG 196
RESULT 8
AAW64591
ID AAW64591 standard; Protein; 1765 AA.
XX
AC AAW64591;
XX
XX 23-OCT-1998 (first entry)
DT Human SRCR protein.
DE
XX Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
KW nervous system; medullo-blastoma; glioma; breast; detection;
KW autoantibody; ss.
XX

OS Homo sapiens.
XX WO9830687-A2.
PN 16-JUL-1998.
XX 09-JAN-1998; 98WO-DE000096.
PF 18-JUL-1997; 97DE-1030997.
PR 09-JAN-1997; 97DE-1000519.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA Mollenhauer J, Poustka A;
XX WPI; 1998-399136/34.
DR N-PSDB; AAV49652.
XX Proteins containing scavenger receptor, cysteine rich domain -
PT useful for diagnosis and treatment of tumours
PT
XX Claim 2; Fig 2; 54pp; German.
XX This sequence represents a human protein which contains a SRCR (scavenger
CC receptor, cysteine-rich) domain. The gene and encoded protein can be used
CC to diagnose or treat tumours, particularly of the nervous system
CC (medullo-blastoma or glioma) or breast. The DNA sequence and probes
CC derived from it, are used to identify genes that express SRCR-domain
CC containing proteins, to determine the form in which these proteins exist
CC and to assess the significance of individual forms on cellular
CC properties. The protein can be used to detect the presence of
CC autoantibodies and antibodies which regulate its expression.
XX
SQ Sequence 1785 AA;
Query Match 52.3%; Score 46; DB 19; Length 1785;
Best Local Similarity 64.3%; Pred. No. 3.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
OY 4 PSNCG--LYSSG 15
Db 1134 PSSNCGGFLFYASG 1147
RESULT 9
ABB05007
ID ABB05007 standard; Protein; 2460 AA.
XX
AC ABB05007;
XX
XX 21-MAR-2002 (first entry)
DT Human reelin protein SEQ ID NO:1.
DE
XX Human; reelin; low density lipoprotein receptor; LDLR; neuroprotective;
KW extracellular glycoprotein; nontropic; anti-lipaeic; Alzheimer's disease;
KW neurodegenerative disorder; neuronal regeneration; cognitive function;
KW lipid metabolism disease; memory; developmental disorder.
XX
OS Homo sapiens.
XX
PN US6323177-B1.
XX
PD 27-NOV-2001.
XX
PF 16-JUN-1999; 99US-0334220.
PR 16-JUN-1999; 99US-0334220.
XX (SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA Curran T, D'Arcangelo G;
XX

DR WPI; 2002-096596/13.
DR N-PSDB; ABA92604.
XX
PT Novel composition useful for screening compounds that modulate Reelin
PT binding to low density lipoprotein receptor, comprising an isolated
PT Reelin polypeptide and low density lipoprotein receptor -
XX
PS Claim 16; Column 31-48; 45pp; English.
XX
CC The present invention describes a composition (I) comprising an
CC isolated reelin protein (II) bound to an isolated low density lipoprotein
CC receptor (LDLR) (III). (II) is an extracellular glycoprotein of
CC approximately 385 kDa containing a small region of similarity with
CC F-spondin at the N terminus, a stretch of positively charged amino
CC acids at the C terminus, and a series of eight internal repeats of
CC 350-390 amino acids, each repeat containing two related sub-domains
CC that flank a pattern of conserved cysteine residues known as an
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,
CC neurotropic and antilipaeamic activities, and can be used as a modulator
CC of reelin-LDLR interaction. (I) is useful in screen for compounds that
CC modulate reelin binding to an LDLR, in an assay system, where the assay
CC system comprises a microplate array and an automated robotic
CC microprocessor controlled system for adding and removing reagents to
CC the microplate array. The compounds identified by the above screening
CC method are useful as therapeutic agents to provide or alleviate a
CC diverse spectrum of diseases including neurodegenerative disorders such
CC as Alzheimer's disease, to facilitate neuronal regeneration after
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance
CC cognitive functions and memory or to ameliorate other developmental
CC disorders. The present sequence represents human reelin, which is used
CC in the exemplification of the present invention.
XX
SQ Sequence 3460 AA;

Query Match 52.3%; Score 46; DB 23; Length 3460;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

CY 3 YPSGNCGLY--YSS 14
||| ||| |||
DB 1973 YPGGNIGLYCPYSS 1986

RESULT 10
ABB05008
ID ABB05008 standard; Protein; 3461 AA.
XX
AC ABB05008;
XX
DT 21-MAR-2002 (first entry)
XX
DE Mouse reelin protein SEQ ID NO:2.
XX
KW Mouse; reelin; low density lipoprotein receptor; LDLR; neuroprotective;
KW extracellular glycoprotein; neurotropic; antilipaeamic; Alzheimer's disease;
KW neurodegenerative disorder; neuronal regeneration; cognitive function;
KW lipid metabolism disease; memory; developmental disorder.
XX
OS Mus musculus.
XX
PN US6323177-B1.
XX
PD 27-NOV-2001.
XX
PF 16-JUN-1999; 99US-0334220.
XX
PR 16-JUN-1999; 99US-0334220.
XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Curran T, D'Arcangelo G;
XX
DR WPI; 2002-096596/13.

DR N-PSDB; ABA92603.
XX
PT Novel composition useful for screening compounds that modulate Reelin
PT binding to low density lipoprotein receptor, comprising an isolated
PT Reelin polypeptide and low density lipoprotein receptor -
XX
PS Claim 16; Column 47-64; 45pp; English.
XX
CC The present invention describes a composition (I) comprising an
CC isolated reelin protein (II) bound to an isolated low density lipoprotein
CC receptor (LDLR) (III). (II) is an extracellular glycoprotein of
CC approximately 385 kDa containing a small region of similarity with
CC F-spondin at the N terminus, a stretch of positively charged amino
CC acids at the C terminus, and a series of eight internal repeats of
CC 350-390 amino acids, each repeat containing two related sub-domains
CC that flank a pattern of conserved cysteine residues known as an
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,
CC neurotropic and antilipaeamic activities, and can be used as a modulator
CC of reelin-LDLR interaction. (I) is useful in screen for compounds that
CC modulate reelin binding to an LDLR, in an assay system, where the assay
CC system comprises a microplate array and an automated robotic
CC microprocessor controlled system for adding and removing reagents to
CC the microplate array. The compounds identified by the above screening
CC method are useful as therapeutic agents to provide or alleviate a
CC diverse spectrum of diseases including neurodegenerative disorders such
CC as Alzheimer's disease, to facilitate neuronal regeneration after
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance
CC cognitive functions and memory or to ameliorate other developmental
CC disorders. The present sequence represents mouse (Mus musculus) reelin,
CC which is used in the exemplification of the present invention.
XX
SQ Sequence 3461 AA;

Query Match 52.3%; Score 46; DB 23; Length 3461;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

CY 3 YPSGNCGLY--YSS 14
||| ||| |||
DB 1974 YPGGNIGLYCPYSS 1987

RESULT 11
ABB57065
ID ABB57065 standard; Protein; 3461 AA.
XX
AC ABB57065;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:129.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99284.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring

PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
PS
XX Claim 2; Page 385-400; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia); by measuring
CC expression levels of particular genes (AB599202 to AB199912, encoding
CC the protein sequences in AB357020 to AB357374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 3461 AA;
Query Match 52.3%; Score 46; DB 23; Length 3461;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 3 YPSGNCGLY--YSS 14
DB 1974 YPGNIGLYCPYSS 1987
RESULT 12
ABG25297
ID ABG25297 standard; Protein; 3470 AA.
XX
AC ABG25297;
XX
DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #25288.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
CS Homo sapiens.
XX
FN WO200175067-A2.
XX
ED 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2002; 2000US-C540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS89484.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 55656; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3470 AA;
Query Match 52.3%; Score 46; DB 22; Length 3470;
Best Local Similarity 71.4%; Pred. No. 6.4e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 3 YPSGNCGLY--YSS 14
DB 1983 YPGNIGLYCPYSS 1996
RESULT 13
AAU80079
ID AAU80079 standard; Protein; 123 AA.
XX
AC AAU80079;
XX
DT 30-JUL-2002 (first entry)
XX
DE Apoptin-associating protein 5 (AAP-5), partial sequence.
XX
KW Cancer; pACT; cytostatic; immunosuppressive; AAP-5;
KW Apoptin-associating protein 5; p53-independent apoptosis;
KW cell proliferation; cell death; autoimmune disease.
XX
OS Homo sapiens.
XX
FN EP1139768-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001EP-0201137.
XX
PR 27-MAR-2000; 2000EP-0201108.
XX
PA (LEAD-) LEADD BV.
XX
PI Noteborn YHM, Rohn JL, Daren-van Oorschot AAAM;
XX
DR NPI; 2001-657960/76.
XX
PT New apoptin-associating protein 5 for inducing p53-independent
PT apoptosis, or for treating cancer or autoimmune disease -
XX
PS Disclosure; Page 23; 44pp; English.
XX
CC The invention relates to an isolated or recombinant nucleic acid
CC which encodes an apoptin-associating protein 5 (AAP-5) capable of
CC providing apoptosis, or its functional fragment. The nucleic acid, a
CC vector comprising the nucleic acid, or a proteinaceous substance is
CC useful for the induction of p53-independent apoptosis. The pharmaceutical
CC composition comprising the nucleic acid, vector comprising the nucleic
CC acid or the proteinaceous substance is also useful for the induction of
CC p53-independent apoptosis, or for the treatment of a disease where

CC enhanced cell proliferation or decreased cell death is observed, such as
CC cancer or autoimmune disease. The present sequence represents the
CC partial amino acid sequence of apoptin-associating protein 5 (AAP-5).
XX

SQ Sequence 123 AA;
Query Match 51.1%; Score 45; DB 22; Length 123;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYSS 14
| | | | |
D5 8 YTSSQCGKYSS 19

RESULT 14
AAU80078
ID AAU80078 standard; Protein; 126 AA.
XX
AC AAU80078;
XX
DT 30-JUL-2002 (first entry)
XX
DE Apoptin-associating protein 5 (AAP-5).
XX
KW Cancer; pACT; cytostatic; immunosuppressive; AAP-5;
KW Apoptin-associating protein 5; p53-independent apoptosis;
KW cell proliferation; cell death; autoimmune disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..3
FT /note= "Encoded by pACT vector"
XX
PN EP1138768-A2.
XX
PD 04-CCT-2001.
XX
PF 27-MAR-2002; 2001EP-0201137.
XX
PR 27-MAR-2000; 2000EP-0201108.
XX
PA (LEAD-) LEADD BV.
XX
PI Noreborn MHM, Rohn JL, Danen-van Corschot AAM;
XX
DR WPI; 2001-657960/76.
DR N-PSDB; ABK50949.
XX
PT New apoptin-associating protein 5 for inducing p53-independent
PT apoptosis, or for treating cancer or autoimmune disease -
PS Claim 10; Fig 2; 44pp; English.
XX
CC The invention relates to an isolated or recombinant nucleic acid
CC which encodes an apoptin-associating protein 5 (AAP-5) capable of
CC providing apoptosis, or its functional fragment. The nucleic acid, a
CC vector comprising the nucleic acid, or a proteinaceous substance is
CC useful for the induction of p53-independent apoptosis. The pharmaceutical
CC composition comprising the nucleic acid, vector comprising the nucleic
CC acid or the proteinaceous substance is also useful for the induction of
CC p53-independent apoptosis, or for the treatment of a disease where
CC enhanced cell proliferation or decreased cell death is observed, such as
CC cancer or autoimmune disease. The present sequence represents the
CC amino acid sequence of apoptin-associating protein 5 (AAP-5).
XX
SQ Sequence 126 AA;
Query Match 51.1%; Score 45; DB 22; Length 126;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYSS 14
| | | | |
D5 11 YTSSQCGKYSS 22

RESULT 15
AAO14806
ID AAO14806 standard; Protein; 126 AA.
XX
AC AAO14806;
XX
DT 08-JUL-2002 (first entry)
XX
DE Human apoptin-associating protein 5, with pACT expression vector region.
XX
KW Human; pACT expression vector; apoptin-associating protein 5; AAP-5;
KW AAP-6; p53-independent apoptosis; gene therapy; cancer;
KW autoimmune disease; apoptosis pathway cascade.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..3
FT /note= "This region is derived from the multiple cloning
FT site in the pACT expression vector"
FT Protein 4..126
FT /note= "This region represents the human apoptin-
FT associating protein 5 (AAP-5)"
XX
PN EP1138765-A1.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2000; 2000EP-0201108.
XX
PR 27-MAR-2000; 2000EP-0201108.
XX
PA (LEAD-) LEADD BV.
XX
PI Noreborn MHM, Rohn JL, Danen-van Corschot AAM;
XX
DR WPI; 2002-012523/C2.
DR N-PSDB; AAL42576.
XX
PT New proteins useful for inducing p53-independent apoptosis and for
PT treating cancer and autoimmune diseases comprises the isolated or
PT recombinant apoptin-associating proteinaceous substance -
PS Disclosure; Fig 3; 28pp; English.
XX
CC The invention comprises the amino acid and coding sequences of two
CC apoptin-associating proteins (AAP-5 and AAP-6). The AAP-5/AAP-6 DNA and
CC protein sequences are useful for inducing p53-independent apoptosis, and
CC the diagnosis/treatment (gene therapy) of cancer and autoimmune diseases.
CC The AAP-5/AAP-6 DNA and protein sequences are also useful for finding
CC additional apoptin-associating proteinaceous substances from the
CC apoptosis pathway cascade. The present amino acid sequence represents the
CC human AAP-5 protein (with a region derived from the multiple cloning
CC site of the pACT expression vector).
XX
SQ Sequence 126 AA;
Query Match 51.1%; Score 45; DB 23; Length 126;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYSS 14
| | | | |
D5 11 YTSSQCGKYSS 22

Search completed: November 5, 2003, 16:46:00

Job time : 3.94053 secs

GenCore version: 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 16:48:06 ; Search time 1.68502 Seconds
(without alignments)
1528.905 Million cell updates/sec

Title: US-09-902-563-13
Perfect score: 88
Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62
Gapcp 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 3
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	88	100.0	15	11	US-09-902-563-18 Sequence 18, Appl
2	88	100.0	15	15	US-10-096-255-18 Sequence 18, Appl
3	88	100.0	432	11	US-09-902-563-4 Sequence 4, Appli
4	88	100.0	432	15	US-10-096-255-4 Sequence 4, Appli
5	88	100.0	439	11	US-09-902-563-2 Sequence 2, Appli
6	88	100.0	439	15	US-10-096-255-2 Sequence 2, Appli
7	45	51.1	126	9	US-09-819-308-2 Sequence 2, Appli
8	45	51.1	210	9	US-09-819-308-10 Sequence 10, Appl
9	41.5	47.2	219	11	US-09-812-355A-24 Sequence 24, Appl
10	41.5	47.2	470	12	US-10-394-557-6 Sequence 6, Appli
11	41.5	47.2	470	12	US-10-223-081-4 Sequence 4, Appli
12	41.5	47.2	470	12	US-10-223-082-4 Sequence 4, Appli
13	41.5	47.2	470	15	US-10-223-085-4 Sequence 4, Appli
14	41.5	47.2	470	15	US-10-223-084-4 Sequence 4, Appli
15	41.5	47.2	470	15	US-10-223-088-4 Sequence 4, Appli

16	41.5	47.2	470	15	US-10-223-090-4 Sequence 4, Appli
17	41.5	47.2	470	15	US-10-223-087-4 Sequence 4, Appli
18	41.5	47.2	470	15	US-10-223-083-4 Sequence 4, Appli
19	41.5	47.2	470	15	US-10-223-089-4 Sequence 4, Appli
20	40	45.5	516	14	US-10-001-870-132 Sequence 132, Appl
21	40	45.5	776	10	US-09-908-193-33 Sequence 33, Appl
22	39.5	44.9	99	11	US-09-229-173-29 Sequence 29, Appl
23	39.5	44.9	212	11	US-09-994-595-150 Sequence 150, Appl
24	39.5	44.9	278	12	US-10-220-511-11 Sequence 11, Appl
25	39.5	44.9	517	12	US-10-032-585-7874 Sequence 7874, Ap
26	39.5	44.9	1391	11	US-09-994-595-8 Sequence 8, Appli
27	39	44.3	588	10	US-09-908-193-34 Sequence 34, Appl
28	39	44.3	673	10	US-09-908-193-46 Sequence 46, Appl
29	39	44.3	695	9	US-09-754-898-191 Sequence 191, App
30	39	44.3	754	10	US-09-908-193-32 Sequence 32, Appl
31	39	44.3	778	10	US-09-908-193-8 Sequence 8, Appli
32	39	44.3	779	10	US-09-908-193-10 Sequence 10, Appl
33	39	44.3	791	10	US-09-908-193-18 Sequence 18, Appl
34	39	44.3	837	11	US-09-946-374-253 Sequence 253, App
35	39	44.3	837	12	US-10-015-387A-253 Sequence 253, App
36	39	44.3	837	12	US-10-006-130A-253 Sequence 253, App
37	39	44.3	837	12	US-10-199-672-454 Sequence 454, App
38	39	44.3	837	12	US-10-241-220-57 Sequence 57, Appl
39	39	44.3	837	12	US-10-006-172A-253 Sequence 253, App
40	39	44.3	837	12	US-10-187-749-454 Sequence 454, App
41	39	44.3	837	12	US-10-194-457-454 Sequence 454, App
42	39	44.3	837	12	US-10-184-642-454 Sequence 454, App
43	39	44.3	837	12	US-10-196-747-454 Sequence 454, App
44	39	44.3	837	12	US-10-015-392A-253 Sequence 253, App
45	39	44.3	837	12	US-10-017-253A-253 Sequence 253, App

ALIGNMENTS

RESULT :
US-09-902-563-18
; Sequence 18, Application US/09902563
; Publication No. US20030099654A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-37
; CURRENT APPLICATION NUMBER: US/09/902,563
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: US 09/442,143
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-563-18

Query Match 100.0%; Score 88; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRYPSGNCGLYSSG 15
| | | | | | | | | | | | | | |
Db 1 DRYPSGNCGLYSSG 15

RESULT 2
US-10-096-255-18
; Sequence 18, Application US/10096255
; Publication No. US20030103974A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-52

```

; CURRENT APPLICATION NUMBER: US/10/096,255
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-096-255-18

Query Match      100.0%; Score 88; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRYPSGNCGLYYSSG 15
      ||| ||||| |||||
DB      1 DRYPSGNCGLYYSSG 15

RESULT 3
US-09-902-563-4
; Sequence 4, Application US/09902563
; Publication No. US20030099654A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-37
; CURRENT APPLICATION NUMBER: US/09/902,563
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: US 09/442,143
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Murine fgl2
US-09-902-563-4

Query Match      100.0%; Score 88; DB 11; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRYPSGNCGLYYSSG 15
      ||||| ||||| |||||
DB      357 DRYPSGNCGLYYSSG 371

RESULT 4
US-10-096-255-4
; Sequence 4, Application US/10096255
; Publication No. US20030103974A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-52
; CURRENT APPLICATION NUMBER: US/10/096,255
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Murine fgl2
US-09-902-563-4
```

```

US-10-096-255-4

Query Match      100.0%; Score 88; DB 15; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRYPSGNCGLYYSSG 15
      ||||| ||||| |||||
DB      357 DRYPSGNCGLYYSSG 371

RESULT 5
US-09-902-563-2
; Sequence 2, Application US/09902563
; Publication No. US20030099654A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-37
; CURRENT APPLICATION NUMBER: US/09/902,563
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: US 09/442,143
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fgl2
US-09-902-563-2

Query Match      100.0%; Score 88; DB 11; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRYPSGNCGLYYSSG 15
      ||||| ||||| |||||
DB      364 DRYPSGNCGLYYSSG 378

RESULT 6
US-10-096-255-2
; Sequence 2, Application US/10096255
; Publication No. US20030103974A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-52
; CURRENT APPLICATION NUMBER: US/10/096,255
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fgl2
US-10-096-255-2

Query Match      100.0%; Score 88; DB 15; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRYPSGNCGLYYSSG 15
      ||||| ||||| |||||
DB      364 DRYPSGNCGLYYSSG 378

RESULT 7
```

US-09-819-308-2
; Sequence 2, Application US/09819308
; Patent No. US20020019040A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020019040Aleborn, Mathieu
; APPLICANT: Danen-van Oorschot, Astrid
; APPLICANT: Rohn, Jennifer
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN
; FILE REFERENCE: 2906-4820US
; CURRENT APPLICATION NUMBER: US/09/819,308
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 126
; TYPE: PRT
; ORGANISM: vector pMT2SM-AAP-5
US-09-819-308-2

Query Match 51.1%; Score 45; DB 9; Length 126;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 3 YPSGNCGLYYSS 14
| | | | |
Db 11 YTSSQCGKYSS 22

RESULT 8
US-09-819-308-10
; Sequence 10, Application US/09819308
; Patent No. US20020019040A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020019040Aleborn, Mathieu
; APPLICANT: Danen-van Oorschot, Astrid
; APPLICANT: Rohn, Jennifer
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN
; FILE REFERENCE: 2906-4820US
; CURRENT APPLICATION NUMBER: US/09/819,308
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 210
; TYPE: PRT
; ORGANISM: open reading frame of AAP-5
US-09-819-308-10

Query Match 51.1%; Score 45; DB 9; Length 210;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 3 YPSGNCGLYYSS 14
| | | | |
Db 95 YTSSQCGKYSS 106

RESULT 9
US-09-832-355A-24
; Sequence 24, Application US/09832355A
; Publication No. US2003027751A1
; GENERAL INFORMATION:
; APPLICANT: Koveshi, Imre
; APPLICANT: Kessler, Paul
; TITLE OF INVENTION: VEGF FUSION PROTEINS
; FILE REFERENCE: 205654
; CURRENT APPLICATION NUMBER: US/09/832,355A
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 219
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-832-355A-24
Query Match 47.2%; Score 41.5; DB 11; Length 219;
Best Local Similarity 60.0%; Pred. No. 93;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1 DRYPSGNCGLYYSSG 15
: | | | | |
Db 153 DSY-SGNCALYQRRG 166
RESULT 10
US-10-394-557-6
; Sequence 6, Application US/10394557
; Publication No. US20030175289A1
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; Gurney, Austin L.
; TITLE OF INVENTION: Tie Ligands
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/394,557
; FILING DATE: 21-Mar-2003
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,821
; FILING DATE: 19-SEPT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1130
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-394-557-6

Query Match 47.2%; Score 41.5; DB 12; Length 470;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYYSSG 15
: | | | | |
Db 404 DSY-SGNCALYQRRG 417

RESULT 11
US-10-223-081-4
; Sequence 4, Application US/10223081
; Publication No. US20030186866A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.

```

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P3235PIC7
; CURRENT APPLICATION NUMBER: US/10/223,081
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/213,637
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-081-4
```

```

Query Match 47.2%; Score 41.5; DB 12; Length 470;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
```

```

QY 1 DRYPSGNCGLYSSG 15
      |||||
Db 404 DSY-SGNCALYCRGG 417
```

```

RESULT 12
US-10-223-082-4
; Sequence 4, Application US/10223082
; Publication No. US20030191059A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P3235PIC3
; CURRENT APPLICATION NUMBER: US/10/223,082
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/213,637
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 09/643,657
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23322
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-082-4
```

```

Query Match 47.2%; Score 41.5; DB 12; Length 470;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
```

```

QY 1 DRYPSGNCGLYSSG 15
      |||||
Db 404 DSY-SGNCALYCRGG 417
```

```

RESULT 13
US-10-223-085-4
; Sequence 4, Application US/10223085
; Publication No. US20030100497A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P3235PIC10
; CURRENT APPLICATION NUMBER: US/10/223,085
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/213,637
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
```

```
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 09/643,657
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-085-4

Query Match          47.2%; Score 41.5; DB 15; Length 470;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15
    | | | | | | |
Db 404 DSY-SGNCALYQRGG 417

RESULT 14
US-10-223-084-4
; Sequence 4, Application US/10223084
; Publication No. US20030105011A;
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P3235P1C5
; CURRENT APPLICATION NUMBER: US/10/223,084
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-085-4
```

```
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-084-4

Query Match          47.2%; Score 41.5; DB 15; Length 470;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15
    | | | | | | |
Db 404 DSY-SGNCALYQRGG 417

RESULT 15
US-10-223-088-4
; Sequence 4, Application US/10223088
; Publication No. US20030105012A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P3235P1C6
; CURRENT APPLICATION NUMBER: US/10/223,088
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 09/643,657
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 393
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-088-4
```

Query Match 47.2%; Score 41.5; DB 15; Length 475;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 DRYPSGNCGJYYSSG 15
| | | | | | | | | |
Db 404 DSY-SGNCALYQRGG 417

Search completed: November 5, 2003, 16:56:34
Job time : 1.68502 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 16:44:23 ; Search time 1.05727 Seconds
(without alignments)
603.285 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	15	4	US-09-442-143A-18
2	88	100.0	432	4	US-09-442-143A-4
3	88	100.0	439	4	US-09-442-143A-2
4	46	52.3	666	4	US-09-341-587-1
5	46	52.3	1785	4	US-09-341-587-3
6	46	52.3	3460	4	US-09-334-220-1
7	46	52.3	3461	4	US-09-334-220-2
8	43	48.9	460	4	US-09-198-452A-7
9	42	47.7	409	4	US-09-328-352-6858
10	41.5	47.2	470	2	US-08-933-821-6
11	41.5	47.2	470	3	US-08-960-507-6
12	41.5	47.2	470	4	US-09-136-828-6
13	41.5	47.2	470	4	US-09-332-928A-6
14	41.5	47.2	470	4	US-09-136-801-6
15	41.5	47.2	470	4	US-09-332-929-6
16	41.5	47.2	470	4	US-09-333-075-6
17	41.5	47.2	470	4	US-09-202-088A-6
18	41.5	47.2	470	4	US-09-333-077-6
19	41	46.6	39	1	US-08-390-882A-1
20	41	46.6	39	1	US-08-390-882A-2
21	40	45.5	1041	1	US-08-494-714-2
22	40	45.5	1041	5	PCT-US96-10782-2
23	39.5	44.9	99	2	US-08-537-400-19
24	39.5	44.9	212	4	US-09-106-568E-150
25	39.5	44.9	1391	4	US-09-106-568E-8
26	39	44.3	38	1	US-08-451-472-9
27	39	44.3	39	1	US-08-390-882A-3

28	39	44.3	78	4	US-09-732-210-338	Sequence 338, Appl
29	39	44.3	79	4	US-09-732-210-343	Sequence 343, Appl
30	39	44.3	79	4	US-09-732-210-347	Sequence 347, Appl
31	39	44.3	80	4	US-09-732-210-346	Sequence 346, Appl
32	39	44.3	81	4	US-09-732-210-337	Sequence 337, Appl
33	39	44.3	524	4	US-09-198-452A-52	Sequence 52, Appl
34	39	44.3	700	4	US-09-408-647A-2	Sequence 2, Appl
35	39	44.3	1077	4	US-09-390-234-12	Sequence 12, Appl
36	39	44.3	2254	2	US-08-286-819A-28	Sequence 28, Appl
37	39	44.3	2254	3	US-08-980-357-28	Sequence 28, Appl
38	38.5	43.8	406	3	US-09-108-020-4	Sequence 4, Appl
39	38.5	43.8	406	3	US-09-108-020-42	Sequence 42, Appl
40	38	43.2	256	4	US-09-252-991A-23974	Sequence 23974, A
41	38	43.2	308	4	US-09-252-991A-27921	Sequence 27921, A
42	38	43.2	438	4	US-09-252-991A-21521	Sequence 21521, A
43	38	43.2	1901	4	US-09-738-946-12	Sequence 12, Appl
44	38	43.2	2465	2	US-08-596-291-3	Sequence 3, Appl
45	38	43.2	2465	3	US-09-100-804-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-442-143A-18

; Sequence 18, Application US/09442143A

; Patent No. 6403089

; GENERAL INFORMATION:

; APPLICANT: Levy, Gary

; APPLICANT: Clark, David A.

; TITLE OF INVENTION: Methods of Modulating Immune Coagulation

; FILE REFERENCE: 9579-14

; CURRENT APPLICATION NUMBER: US/09/442,143A

; CURRENT FILING DATE: 1999-11-15

; PRIOR APPLICATION NUMBER: US 60/046,537

; PRIOR FILING DATE: 1997-05-17

; PRIOR APPLICATION NUMBER: US 60/061,684

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 18

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-442-143A-18

Query Match 100.0%; Score 88; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRYPSGNCGLYSSG 15

Db 1 DRYPSGNCGLYSSG 15

RESULT 2

US-09-442-143A-4

; Sequence 4, Application US/09442143A

; Patent No. 6403089

; GENERAL INFORMATION:

; APPLICANT: Levy, Gary

; APPLICANT: Clark, David A.

; TITLE OF INVENTION: Methods of Modulating Immune Coagulation

; FILE REFERENCE: 9579-14

; CURRENT APPLICATION NUMBER: US/09/442,143A

; CURRENT FILING DATE: 1999-11-15

; PRIOR APPLICATION NUMBER: US 60/046,537

; PRIOR FILING DATE: 1997-05-17

; PRIOR APPLICATION NUMBER: US 60/061,684

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

```
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Mus mus fgl2
US-09-442-143A-4

Query Match      100.0%; Score 88; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRYPSGNCGLYSSG 15
      |||||
Db      357 DRYPSGNCGLYSSG 371

RESULT 3
US-09-442-143A-2
; Sequence 2, Application: US/09442:143A
; Patent No. 6403089
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-14
; CURRENT APPLICATION NUMBER: US/09/442,143A
; CURRENT FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fgl2
US-09-442-143A-2

Query Match      100.0%; Score 88; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DRYPSGNCGLYSSG 15
      |||||
Db      364 DRYPSGNCGLYSSG 378

RESULT 4
US-09-341-587-1
; Sequence 1, Application: US/09341:587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-587-1

Query Match      52.3%; Score 46; DB 4; Length 666;
Best Local Similarity 64.3%; Pred. No. 36;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      4 PSNGCG--LYYSSG 15
      |||||
Db      183 PSSNCGGFLFYASG 196
```

```
RESULT 5
US-09-341-587-3
; Sequence 3, Application: US/09341:587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1785
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-587-3

Query Match      52.3%; Score 46; DB 4; Length 1785;
Best Local Similarity 64.3%; Pred. No. 98;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY      4 PSNGCG--LYYSSG 15
      |||||
Db      1134 PSSNCGGFLFYASG 1147

RESULT 6
US-09-334-220-1
; Sequence 1, Application: US/09334:220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-334-220-1

Query Match      52.3%; Score 46; DB 4; Length 3460;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY      3 YPSGNCGLY--YSS 14
      |||||
Db      1973 YPGNIGLYCPYSS 1986

RESULT 7
US-09-334-220-2
; Sequence 2, Application: US/09334:220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
```

```
(
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3461
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-334-220-2

Query Match 52.3%; Score 46; DB 4; Length 3461;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNCGLY--YSS 14
: ||| ||| |||
Db 1974 YPGNIGLYCPYSS 1987

RESULT 8
US-09-198-452A-7
; Sequence 7, Application: US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; thereof and uses thereof, in particular for the diagnosis, prevd
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 7
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-7

Query Match 48.9%; Score 43; DB 4; Length 460;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 3 YPSGNCGLY 12
: ||| |||
Db 7 FPGNCNCYY 16

RESULT 9
US-09-328-352-6858
; Sequence 6858, Application: US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6858
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6858

Query Match 47.7%; Score 42; DB 4; Length 409;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SGNCGLY 11
: |||||
Db 396 SGNCGLY 402
```

```
RESULT 10
US-08-933-821-6
; Sequence 6, Application: US/08933821
; Patent No. 5972338
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Tie Ligands
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,821
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1130
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-933-821-6

Query Match 47.2%; Score 41.5; DB 2; Length 470;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15
: ||| ||| |||
Db 404 DSY-SGNCALYQRGG 417

RESULT 11
US-08-960-507-6
; Sequence 6, Application: US/08960507
; Patent No. 6057435
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Tie Ligands
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,507
```

```

; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1130p1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-960-507-6

Query Match 47.2%; Score 41.5; DB 3; Length 470;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15
| | | | | | | | |
Db 404 DSY-SGNCALYQRG 417

RESULT 12
US-09-136-828-6
; Sequence 6, Application US/09136828
; Patent No. 6350450
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Tie Ligands
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,828
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1130r1a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-09-136-828-6

Query Match 47.2%; Score 41.5; DB 4; Length 470;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15
| | | | | | | | |
Db 404 DSY-SGNCALYQRG 417
```

```

; FILING DATE:
; CLASSIFICATION:
; US-09-332-928A-6
; Sequence 6, Application US/09332928A
; Patent No. 6368853
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: Tie Ligands
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/332,928A
; FILING DATE: 14-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/933,821
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1130
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
;
US-09-332-928A-6

Query Match 47.2%; Score 41.5; DB 4; Length 470;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15
| | | | | | | | |
Db 404 DSY-SGNCALYQRG 417

RESULT 14
US-09-136-801-6
; Sequence 6, Application US/09136801
; Patent No. 6413770
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Botstein, David
; APPLICANT: Goddard, Audrey
; APPLICANT: Roy, Margaret
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Tumas, Daniel
; APPLICANT: Schwall, Ralph
; TITLE OF INVENTION: Tie Ligand Homologues
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
```

COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,801
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1130P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 470 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-136-801-6

Query Match 47.2%; Score 41.5; DB 4; Length 470;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYYSSG 15
|!|||||||
Db 404 DSY-SGNCALYQRGG 417

RESULT 15
US-09-332-929-6
Sequence 6, Application US/09332929
Patent No. 6420542
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
TITLE OF INVENTION: Tie Ligands
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/332,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/933,821
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1130
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 470 amino acids
TYPE: Amino Acid

TOPOLOGY: Linear
US-09-332-929-6
Query Match 47.2%; Score 41.5; DB 4; Length 470;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1 DRYPSGNCGLYYSSG 15
|!|||||||
Db 404 DSY-SGNCALYQRGG 417
Search completed: November 5, 2003, 16:49:00
Job time : 2.05727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: November 5, 2003, 16:55:36 ; Search time 20 seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPGNCGLYSSG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168652 residues

Total number of hits satisfying chosen parameters: 2520

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match: 0%
Maximum Match: 100%
Listing first 45 summaries

Database : PIR 75:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	31	35.2	14	2 A58963	alpha-conotoxin Cn
2	31	35.2	15	1 NTKNAG	alpha-conotoxin GI
3	26	29.5	15	2 PQ0017	terminal protein -
4	24	27.3	10	2 C41946	T-cell receptor ga
5	23.5	26.7	12	2 S25056	Ig heavy chain - m
6	23	26.1	10	2 PT0243	Ig heavy chain CRD
7	23	26.1	13	1 JTJG3	tremorogen a-13 -
8	23	26.1	14	2 PH0792	T-cell receptor al
9	23	26.1	15	2 PH1366	Ig heavy chain DJ
10	23	26.1	15	2 PH0750	T-cell receptor be
11	23	26.1	15	2 PJ0110	complement factor
12	22.5	25.6	11	2 B41946	T-cell receptor ga
13	22	25.0	9	2 S19329	sperm-activating p
14	22	25.0	11	2 PT0302	Ig heavy chain CRD
15	22	25.0	13	1 NTKN2G	alpha-conotoxin GI
16	22	25.0	13	2 I50173	alpha-2 collagen -
17	22	25.0	14	1 NTKN1M	alpha-conotoxin MI
18	22	25.0	14	2 PH1305	Ig heavy chain DJ
19	22	25.0	15	2 C37765	hypothetical prote
20	21	23.9	9	2 PT0288	Ig heavy chain CRD
21	21	23.9	9	2 G41946	T-cell receptor ga
22	21	23.9	9	2 PQ0027	pev-tachykinin - p
23	21	23.9	10	2 B38887	T-cell receptor ga
24	21	23.9	11	2 PH1584	Ig H chain V-D-J r
25	21	23.9	12	2 H41946	T-cell receptor ga
26	21	23.9	13	2 A28953	alpha-conotoxin SI
27	21	23.9	13	2 A61210	antibiotic GE2270
28	21	23.9	15	2 PT0085	protein QA60027 -
29	21	23.9	15	2 S47387	T-cell antigen rec

30	21	23.9	15	2 S03955	acidic fibroblast
31	21	23.9	15	2 S08301	epidermal growth f
32	20	22.7	6	2 PT0652	T-cell receptor be
33	20	22.7	11	2 S68637	acetylcholinestera
34	20	22.7	13	2 PH1596	Ig H chain V-D-J r
35	20	22.7	14	2 S58862	botulinum neurotox
36	20	22.7	14	2 S58866	botulinum neurotox
37	20	22.7	14	2 PT0252	Ig heavy chain CRD
38	20	22.7	14	2 PH1597	Ig H chain V-D-J r
39	19	21.6	7	2 A11483	aspartate transami
40	19	21.6	10	2 A31571	hypertrehalosemic/
41	19	21.6	12	2 S09082	proteasome chain 1
42	19	21.6	12	2 PH0802	T-cell receptor al
43	19	21.6	13	2 PH0788	T-cell receptor al
44	19	21.6	14	2 PA0109	porin por 1B - Ara
45	19	21.6	14	2 PA0045	porin por1 - Arabi

ALIGNMENTS

RESULT 1

A58963
alpha-conotoxin CnIA - cone shell (Conus consors)
N:Contains: alpha-conotoxin CnIB
C:Species: Conus consors
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: A58963
R:Favreau, F.; Krimm, I.; Le Gall, F.; Bobenrieth, M.C.; Lamthanh, H.; Bouet, F.; Serv
Biochemistry 38, 6317-6326, 1999
A:Title: Biochemical characterization and nuclear magnetic resonance structure of nove
A:Reference number: A58963; MUID:99255390; PMID:10320362
A:Accession: A58963
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-14 <FAV>
C:Superfamily: alpha-conotoxin
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neur
E:1-14/Product: alpha-conotoxin CnIA #status experimental <MATA>
F:3-14/Product: alpha-conotoxin CnIB #status experimental <MATA>
F:3-8,4-14/Disulfide bonds: #status experimental
F:14/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 35.2%; Score 31; DB 2; Length 14;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLYYS 13

|||

Db 8 CGKYYS 13

RESULT 2

NTKNAG

alpha-conotoxin GIA (validated) - cone shell (Conus geographus)
N:Alternate names: alpha-CTX-GIA
N:Contains: alpha-conotoxin GI
C:Species: Conus geographus (geography cone)
C:Date: 24-Sep-1981 #sequence_revision 24-Sep-1981 #text_change 15-Sep-2000
C:Accession: A01782
R:Gray, W.R.; Luque, A.; Olivera, B.M.; Barrett, J.; Cruz, L.J.
J. Biol. Chem. 256, 4734-4740, 1981
A:Title: Peptide toxins from Conus geographus venom.
A:Reference number: A92320; MUID:81191854; PMID:7014556
A:Accession: A01782
A:Molecule type: protein
A:Residues: 1-15 <GRA>
R:Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M.
J. Biol. Chem. 258, 12247-12251, 1983
A:Title: Conotoxin MI. Disulfide bonding and conformational states.
A:Reference number: A92396; MUID:84032400; PMID:6630187
A:Contents: annotation; disulfide bonds
R:Guddat, L.W.; Shan, L.; Martin, J.L.; Edmundson, A.B.; Gray, W.R.

```
submitted to the Brookhaven Protein Data Bank, May 1996
A;Reference number: A66253; PDB:1NCT
A;Contents: annotation; X-ray crystallography, 1.2 angstroms, residues 1-13
R;Gudwat, L.W.; Martin, C.A.; Shan, L.; Edmundson, A.B.; Gray, W.R.
Biochemistry 35, 11329-11335, 1996
A;Title: Three-dimensional structure of the alpha-conotoxin GI at 1.2 angstroms resolution
A;Reference number: A58592; MUID:96378624; PMID:8784187
A;Contents: annotation; X-ray crystallography, 1.2 angstroms
R;Pardi, A.; Galdes, A.; Florance, J.; Manicasse, D.
Biochemistry 28, 5494-5501, 1989
A;Title: Solution structures of alpha-conotoxin GI determined by two-dimensional NMR spectroscopy
A;Reference number: A30629; MUID:89375269; PMID:2775719
A;Contents: annotation; conformation by (1)H-NMR
C;Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynaptic antagonist
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; blocked carboxyl end
F;1-15/Product: conotoxin GIA #status experimental <GIA>
F;1-13/Product: conotoxin GI #status experimental <GIC>
F;2-7,3-13/Disulfide bonds: #link GIA #status predicted
F;2-7,3-13/Disulfide bonds: #link GIC #status experimental
F;13/Modified site: amidated carboxyl end (Cys) (amide in mature form from following glycosylation)
F;15/Modified site: blocked carboxyl end (Lys) (probably amidated) #status experimental

Query Match          35.2%; Score 31; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      8 CGLYYSSG 15
      ||| |||
Db      7 CGRHYSCG 14

RESULT 3
PQ0017
C;Terminal protein - phage M2 (fragment)
C;Species: phage X2
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: PQ0017
R;Matsumoto, K.; Takano, H.; Kim, C.I.; Hirokawa, H.
Gene 84, 247-255, 1989
A;Title: Primary structure of bacteriophage X2 DNA polymerase: conserved segments within the polymerase gene
A;Reference number: CQ0161; MUID:90128268; PMID:2515115
A;Accession: PQ0017
A;Molecule type: DNA
A;Residues: 1-15 <MAT>
A;Cross-references: GB:M33144; NID:G215507; PIDN:AAA32367.1; PID:G215503
C;Genetics:
A;Gene: E
C;Superfamily: phage PZA terminal protein

Query Match          29.5%; Score 26; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 DRYPSGNCGL 10
      ||| |
Db      1 DRYERGDNVL 10

RESULT 4
C41946
T-cell receptor gamma chain (1t.60) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1995
C;Accession: C41946
R;Whetseil, M.; Mosley, R.L.; Whetseil, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma gene
A;Reference number: A41946; MUID:92049316; PMID:1658619
A;Accession: C41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-10 <WHE>
```

```
C;Keywords: T-cell receptor

Query Match          27.3%; Score 24; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      8 CGLYYSSG 15
      ||| |||
Db      2 CAVWSSSG 9

RESULT 5
S25056
Ig heavy chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 25-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 20-Jun-2000
C;Accession: S25056
R;Jacob, J.; Kelsoe, G.
submitted to the EMBL Data Library, July 1992
A;Description: In situ studies on the primary immune response to (4-hydroxy-3-nitrophenyl)
A;Reference number: S25024
A;Accession: S25056
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-12 <JAC>
A;Cross-references: EMBL:X67386; NID:G50927; PIDN:CAA47798.1; PID:gl333920
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match          26.7%; Score 23.5; DB 2; Length 12;
Best Local Similarity 46.2%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY      2 RYPSGNCGLYYSS 14
      ||| |||
Db      2 RYP-----YYGS 9

RESULT 6
PT0243
Ig heavy chain CDR3 region (clone 2-103A) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0243
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.L.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0243
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match          26.1%; Score 23; DB 2; Length 10;
Best Local Similarity 71.4%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      9 CGLYYSSG 15
      ||| |||
Db      2 GYVDSSG 8

RESULT 7
JTCG3
tremorogen a-13 - jelly fungus (Tremella mesenterica)
C;Species: Tremella mesenterica
C;Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 07-Nov-1997
C;Accession: A01641
R;Sakagami, Y.; Yoshida, M.; Isogai, A.; Suzuki, A.
Science 212, 1525-1527, 1981
A;Title: Peptide sex hormones inducing conjugation tube formation in compatible mating
A;Reference number: A94256
```

```
A;Accession: A01641
A;Molecule type: protein
A;Residues: 1-13 <SAK>
C;Comment: Tremmerogen a-13 is produced by the a mating-type cells and induces formation
C;Superfamily: tremmerogen a-13
C;Keywords: extracellular protein; hormone; lipoprotein; pheromone; prenylated cysteine
F;13/Binding site: farnesy (Cys) (covalent) #status experimental
F;13/Modified site: methyl ester carboxy; end (Cys) #status absent

Query Match      26.1%; Score 23; DB 1; Length 13;
Best Local Similarity 80.6%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 3;

QY      4 PSGNC 8
DB      9 PSGVC 13

RESULT 8
PH0792
T-cell receptor alpha chain (H2 V-alpha-2.7A19) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0792
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0792
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60897
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match      26.1%; Score 23; DB 2; Length 14;
Best Local Similarity 45.5%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      5 SGNGGLYYSSG 15
DB      4 SANSQTYQRF 14

RESULT 9
PH1366
IG heavy chain DJ region (clone C11-106) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1366
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PH1302; MUID:93094761; PMID:1460419
A;Accession: PH1366
A;Molecule type: DNA
A;Residues: 1-15 <WAS>
C;Keywords: heterotetramer; immunoglobulin

Query Match      26.1%; Score 23; DB 2; Length 15;
Best Local Similarity 45.5%; Pred. No. 2.2e+03;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      5 SGNGGLYYSSG 15
DB      4 SGVYGDYWGQG 14

RESULT 10
PH0750
T-cell receptor beta chain (C11) - mouse (fragment)
C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0750
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0750
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X06841
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match      26.1%; Score 23; DB 2; Length 15;
Best Local Similarity 44.4%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      4 PSGNCGLY 12
DB      5 PTANTGQLY 13

RESULT 11
PL0110
complement factor B1-Bb and B2-Bb - guinea pig (fragment)
C;Species: Cavia porcellus (guinea pig)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 15-Nov-1996
C;Accession: PL0110
R;Matsushita, M.; Okada, H.
Mol. Immunol. 26, 669-676, 1989
A;Title: Two forms of guinea pig factor B of the alternative complement pathway with d
A;Reference number: A93136; MUID:89384686; PMID:2779589
A;Accession: PL0110
A;Molecule type: protein
A;Residues: 1-15 <MAT>
C;Keywords: complement alternate pathway; glycoprotein

Query Match      26.1%; Score 23; DB 2; Length 15;
Best Local Similarity 37.5%; Pred. No. 2.2e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      4 PSGNCGLY 11
DB      6 PAGSKNIY 13

RESULT 12
B41946
T-cell receptor gamma chain (it.57) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: B41946
R;Whetsell, V.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma
A;Reference number: A41946; MUID:92049316; PMID:1658619
A;Accession: B41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-11 <WHE>
C;Keywords: T-cell receptor

Query Match      25.6%; Score 22.5; DB 2; Length 11;
Best Local Similarity 55.6%; Pred. No. 2e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY      8 CGLY-YSSG 15
DB      2 CAVWRYSYG 10

RESULT 13
```

S19329
sperm-activating peptide SAP - sea urchin (Stomopneustes variolus)
C;Species: Stomopneustes variolus
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Aug-2000
C;Accession: S19329
R;Yoshino, K.; Takao, T.; Shimonishi, Y.; Suzuki, N.
FEBS Lett. 294, 179-182, 1991
A;Title: Determination of the amino acid sequence of an intramolecular disulfide linkage
A;Reference number: S19329; MUID:92097763; PMID:1756859
A;Accession: S19329
A;Molecule type: protein
A;Residues: 1-9 <YOS>
C;Superfamily: unassigned animal peptides
F;3-8/Disulfide bonds: #status predicted

Query Match 25.0%; Score 22; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSGNC 8
| | |
Db 4 PEGKC 8

RESULT 14
PT0302
Ig heavy chain CRD3 region (clone 5-112) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0302
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0302
A;Molecule type: DNA
A;Residues: 1-11 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 25.0%; Score 22; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 2.4e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 GLYSSG 15
| | | | |
Db 3 GFRSSG 9

RESULT 15
NTKN2G
alpha-conotoxin GII - cone shell (Conus geographus)
C;Species: Conus geographus (geography cone)
C;Date: 24-Sep-1981 #sequence_revision 24-Sep-1981 #text_change 23-May-1997
C;Accession: A01783
R;Gray, W.R.; Lucue, A.; Olivera, B.M.; Barrett, J.; Cruz, L.J.
J. Biol. Chem. 256, 4734-4740, 1981
A;Title: Peptide toxins from Conus geographus venom.
A;Reference number: A92396; MUID:84032400; PMID:6630187
A;Accession: A01783
A;Molecule type: protein
A;Residues: 1-13 <GRA>
R;Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M.
J. Biol. Chem. 258, 12247-12251, 1983
A;Title: Conotoxin MI. Disulfide bonding and conformational states.
A;Reference number: A92396; MUID:84032400; PMID:6630187
A;Contents: annotation; disulfide bonds
C;Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynaptic
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; blocked carboxyl end; postsynaptic neurotoxin
F;2-7,3-13/Disulfide bonds: #status predicted
F;13/Modified site: blocked carboxyl end (Cys) (probably amidated) #status experimental

Query Match 25.0%; Score 22; DB 1; Length 13;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLYYS 13
| | | | |
Db 7 CGKHFS 12

Search completed: November 5, 2003, 16:59:09
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 16:49:10 : Search time 11 seconds
(without alignments)
64,127 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPSGNCLYSSG 15

Scoring table: BL0SCM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 795

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	31	35.2	14	1	CXAI_CONCN
2	31	35.2	15	1	CXAI_CONGE
3	26	29.5	15	1	TERM_BPM2
4	23	26.1	13	1	TAL3_TREME
5	23	26.1	14	1	MARI_ALTSP
6	23	26.1	15	1	UC14_MAIZE
7	22	25.0	9	1	SAP_STOVA
8	22	25.0	13	1	CXA2_CONGE
9	22	25.0	14	1	CXA1_CONMA
10	21	23.9	9	1	TRP4_LEUMA
11	21	23.9	15	1	CLQA_RAT
12	21	23.9	15	1	FGF1_CANFA
13	20	22.7	10	1	PPCK_FASHE
14	19	21.6	9	1	TKL1_LOCMI
15	19	21.6	10	1	H7P_HELZE
16	19	21.6	14	1	SKK3_LEICU
17	19	21.6	15	1	TAL_TREBR
18	19	21.6	15	1	UC06_MAIZE
19	18	20.5	10	1	AMPN_HELAM
20	18	20.5	13	1	UHAI_HUMAN
21	18	20.5	14	1	CAT2_FASHE
22	18	20.5	14	1	HY14_PIG
23	16	20.5	14	1	LPER_BACLI
24	18	20.5	15	1	KLOM_LUMTE
25	18	20.5	15	1	UC19_MAIZE
26	17	19.3	8	1	ALL1_CYDPO
27	17	19.3	8	1	LCK8_LEJMA
28	17	19.3	9	1	OXYA_SQJAC
29	17	19.3	9	1	CXYT_RABIT
30	17	19.3	10	1	FAR5_MACRS
31	17	19.3	12	1	CX13_CONMR
32	17	19.3	14	1	SODN_STRGR
33	17	19.3	15	1	OBPA_MAMBR

34	16	18.2	9	1	ALL6_CYDPO	P82157 cydia pomon
35	16	18.2	9	1	COW_CONVE	P83047 conus ventr
36	16	18.2	10	1	COXO_THUOB	P80982 thunnus obe
37	16	18.2	10	1	CU30_LOCMI	P11735 locusta mig
38	16	18.2	11	1	CORZ_PERAM	P11496 periplaneta
39	16	18.2	11	1	EFG_CLOPA	P81350 clostridium
40	16	18.2	12	1	FARI_CALVO	P41869 calliphora
41	16	18.2	12	1	FIFI_SARBU	P83349 sarcophaga
42	16	18.2	12	1	FREI_LITIN	P82021 litoria inf
43	16	18.2	12	1	LICH_BACLI	P82907 bacillus li
44	16	18.2	12	1	RFL_CONSP	P58805 conus spuri
45	16	18.2	13	1	LPAA_PORGI	P81411 porphyromon

ALIGNMENTS

RESULT 1
CXAI_CONCN STANDARD; PRT; 14 AA.
AC P56973;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-conotoxin CN1A [Contains: Alpha-conotoxin CN1B].
CS Conus consors (Singed cone).
CC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
CC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
CC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=101297;
RN 1;
RP SEQUENCE, SYNTHESIS, CHARACTERIZATION, AND STRUCTURE BY NMR.
RC TISSUE=Venom;
RX MEDLINE=9255390; PubMed=10320362;
RA Favreau P., Krimm I., le Gall F., Bobenrieth M.J., Lamthanh H.,
Bouet F., Servent D., Molgo J., Menez A., Letourneux Y.,
RA Lancelin J.-M.;
RT "Biochemical characterization and nuclear magnetic resonance
structure of novel alpha-conotoxins isolated from the venom of Conus
consors.";
RL Biochemistry 38:6317-6326(1999).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
INHIBIT THEM. THIS PEPTIDE SEEMS TO BE A POTENT AND SELECTIVE
BLOCKER OF MUSCULAR SUBTYPE OF NACHR.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
FAMILY.
PIR; A58963; A58963.
PDB; 1B45; 09-JUL-99.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation; 3D-structure.
FT PEPTIDE 1 14 ALPHA-CONOTOXIN CN1A.
FT PEPTIDE 3 14 ALPHA-CONOTOXIN CN1B.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14 AMIDATION.
FT HELIX 6 8
FT TURN 9 10
SQ SEQUENCE 14 AA; 1548 MW; DEE91969BF5ESBD CRC64;

Query Match 35.2% Score 31; DB 1; Length 14;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 8 CGLYYS 13
DB 8 CGKYS 13

RESULT 2
CXAI_CONGE

```
AD CXAL CONGE STANDARD; PRT; 15 AA.
AC P01519;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-conotoxin GIA [Contains: Alpha-conotoxin GI (GI)].
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE.
RX MEDLINE=81191854; PubMed=7014556;
RX Gray W.R., Luque F.A., Olivera B.M., Barrett J., Cruz L.J.;
RT "Peptide toxins from Conus geographus venom.";
RL J. Biol. Chem. 256:4734-4740(1981).
RN [2]
RP DISULFIDE BONDS OF GI, AND SYNTHESIS OF GI.
RX MEDLINE=83105694; PubMed=7152021;
RA Nishiuchi Y., Sakakibara S.;
RT "Primary and secondary structure of conotoxin GI, a neurotoxic
RT tridecapeptide from a marine snail.";
RL FEBS Lett. 148:260-262(1982).
RN [3]
RP DISULFIDE BONDS OF GI, AND SYNTHESIS OF GI.
RX MEDLINE=84280842; PubMed=6466616;
RX Gray W.R., Luque F.A., Galyean R., Atherton E., Sheppard R.C.,
RA Stone B.L., Reyes A., Alford J., McIntosh M., Olivera B.M.,
RA Cruz L.J., Rivier J.;
RT "Conotoxin GI: disulfide bridges, synthesis, and preparation of
RT iodinated derivatives.";
RL Biochemistry 23:2796-2802(1984).
RN [4]
RP COMPARISON WITH ALPHA-CONOTOXIN SI AND ALPHA-CONOTOXIN MI.
RX MEDLINE=95034849; PubMed=7947815;
RA Hann R.M., Pagan O.R., Eterovic V.A.;
RT "The alpha-conotoxins GI and MI distinguish between the nicotinic
RT acetylcholine receptor agonist sites while SI does not.";
RL Biochemistry 33:14058-14063(1994).
RN [5]
RP PHARMACOLOGICAL CHARACTERIZATION ON MOUSE MUSCLE-DERIVED BC3H-1 CELLS
RP AND TORPEDO ELECTRIC ORGAN.
RX MEDLINE=95349531; PubMed=7623764;
RA Groebe D.R., Dunn G.M., Levitan E.S., Abramson S.N.;
RT "alpha-Conotoxins selectively inhibit one of the two acetylcholine
RT binding sites of nicotinic receptors.";
RL Mol. Pharmacol. 48:105-111(1995).
RN [6]
RP MUTAGENESIS OF ARG-9.
RX MEDLINE=97317090; PubMed=9174364;
RA Groebe D.R., Gray W.R., Abramson S.N.;
RT "Determinants involved in the affinity of alpha-conotoxins GI and SI
RT for the muscle subtype of nicotinic acetylcholine receptors.";
RL Biochemistry 36:6459-6474(1997).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS) OF GI.
RX MEDLINE=96378624; PubMed=8784187;
RA Guddat L.W., Marcin J.A., Shan L., Edmundson A.R., Gray W.R.;
RT "Three-dimensional structure of the alpha-conotoxin GI at 1.2-A
RT resolution.";
RL Biochemistry 35:11329-11335(1996).
RN [8]
RP STRUCTURE BY NMR OF GI.
RX MEDLINE=89325262; PubMed=2765514;
RA Kobayashi Y., Ohkubo T., Kyogoku Y., Nishiuchi Y., Sakakibara S.,
RA Braun W., Go N.;
RT "Solution conformation of conotoxin GI determined by 1H nuclear
RT magnetic resonance spectroscopy and distance geometry calculations.";
RL Biochemistry 28:4853-4860(1989).
RN [9]
RP STRUCTURE BY NMR OF GI.
RX MEDLINE=89375269; PubMed=2775719;
RA Pardi A., Galdes A., Florence J., Manicote D.;
RT "Solution structures of alpha-conotoxin GI determined by two-
RT dimensional NMR spectroscopy.";
RL Biochemistry 28:5494-5501(1989).
RN [10]
RP STRUCTURE BY NMR OF GI.
RX MEDLINE=98321613; PubMed=9660176;
RA Maslennikov I.V., Sobol A.G., Gladky K.V., Lugovskoy A.A.,
RA Ostrovsky A.G., Tsetlin V.I., Ivanov V.T., Arseniev A.S.;
RT "Two distinct structures of alpha-conotoxin GI in aqueous solution.";
RL Eur. J. Biochem. 254:238-247(1998).
RN [11]
RP STRUCTURE BY NMR OF GI.
RX MEDLINE=98239743; PubMed=9571060;
RA Gehrman J., Alewood P.F., Craik D.J.;
RT "Structure determination of the three disulfide bond isomers of
RT alpha-conotoxin GI: a model for the role of disulfide bonds in
RT structural stability.";
RL J. Mol. Biol. 278:401-415(1998).
RN [12]
RP STRUCTURE BY NMR OF AN ANTITOXIC ANALOG OF GI.
RX MEDLINE=99438341; PubMed=10508392;
RA Mok K.H., Han K.H.;
RT "NMR solution conformation of an antitoxic analogue of alpha-conotoxin
RT GI: identification of a common nicotinic acetylcholine receptor
RT alpha(1)-subunit binding surface for small ligands and alpha-
RT conotoxins.";
RL Biochemistry 38:11895-11904(1999).
CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
CC bind to the nicotinic acetylcholine receptors (nAChR) and thus
CC inhibit them. The higher affinity site for alpha-conotoxin GI is
CC the alpha/delta site on mouse muscle-derived BC3H-1 receptor, and
CC the other site (alpha/gamma site) on nicotinic receptors from
CC Torpedo californica electric organ.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.
DR PIR; A01782; NTKNAG.
DR PDB; INOT; 07-DEC-96.
DR PDB; 1XGA; 16-FEB-99.
DR PDB; 1XGB; 16-FEB-99.
DR PDB; 1XGC; 23-MAR-99.
DR PDB; 1QS3; 06-OCT-99.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation; 3D-structure.
FT PEPTIDE 1 15 ALPHA-CONOTOXIN GIA.
FT PEPTIDE 1 13 ALPHA-CONOTOXIN GI.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD_RES 13 13
FT
FT AMIDATION (G-14 PROVIDE AMIDE GROUP) (IN
FT ALPHA-CONOTOXIN GI).
FT AMIDATION (IN ALPHA-CONOTOXIN GIA).
FT R->A: REDUCTION IN AFFINITY FOR BOTH
FT ALPHA/DELTA AND ALPHA/GAMMA SITES ON
FT BC3H-1 RECEPTORS AND LOSS OF AFFINITY FOR
FT BOTH ALPHA/DELTA AND ALPHA/GAMMA SITES ON
FT TORPEDO RECEPTORS (IN GI).
FT
FT STRAND 2 2
FT HELIX 5 10
FT STRAND 12 12
SQ SEQUENCE 15 AA; 1628 MW; 2AE73EE90F8C2E19 CRC64;
Query Match 35.2%; Score 32; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 8 CGLYSSG 15
||:|:|
Db 7 CGRHSCG 14
RESULT 3
```

TERM_BP22
ID TERM_BP22 STANDARD; PRT; 15 AA.
AC P19897;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA terminal protein (Protein GPI) (Fragment).
GN 3 OR E.
OS Bacteriophage M2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC phi-29-like viruses.
OX NCBI_TaxID=10751;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90128268; PubMed=2515115;
RA Matsumoto K., Takano H., Kim C.I., Hirokawa H.;
RT "Primary structure of bacteriophage M2 DNA polymerase: conserved
segments within protein-priming DNA polymerases and DNA polymerase I
of Escherichia coli.";
RL Gene 84:247-255(1989).
CC -!- FUNCTION: DNA TERMINAL PROTEIN IS LINKED TO THE 5' ENDS OF
BOTH STRANDS OF THE GENOME THROUGH A PHOSPHODIESTER BOND BETWEEN
THE BETA-HYDROXYL GROUP OF A SERINE RESIDUE AND THE 5'-PHOSPHATE
OF THE TERMINAL DEOXYADENYLATE. THIS PROTEIN IS ESSENTIAL FOR DNA
REPLICATION AND IS INVOLVED IN THE PRIMING OF DNA ELONGATION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: M33144; AAA32367.1; -.
DR PIR: PQ0017; PQ0017.
KW Early protein; DNA replication; DNA priming;
KW Covalent protein-DNA linkage.
FT NON TER 1
FT SITE 5 7 CELL ATTACHMENT SITE (POTENTIAL).
FT SEQUENCE 15 AA; 1797 MW; D3CBFAFF8759DEA36 CRC64;

Query Match 29.5%; Score 26; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 DRYPSGNCGL 10
Db 1 DRYERGQVNL 10

RESULT 4
TA13 TREME STANDARD; PRT; 13 AA.
AC P01370;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE Tremetogen A-13.
OS Tremella mesenterica (Celly fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremelliales; Tremellaceae; Tremella.
OX NCBI_TaxID=5217;
RN [1]
RP SEQUENCE.
RA Sakagami Y., Yoshida M., Isogai A., Suzuki A.;
RT "Peptide sex hormones inducing conjugation tube formation in
compatible mating-type cells of Tremella mesenterica.";
RL Science 212:1525-1527(1981).
CC -!- FUNCTION: TREMEROGEN A-13 IS PRODUCED BY THE A MATING-TYPE CELLS
AND INDUCES FORMATION OF CONJUGATION TUBES IN A MATING-TYPE CELLS.
CC PIR: A01641; JTJG3.
KW Lipoprotein; Prenylation; Pheromone.

FT LIPID 13 13 FARNESYL.
SQ SEQUENCE 13 AA; 1204 MW; 680304A9697BA864 CRC64;

Query Match 26.1%; Score 23; DB 1; Length 13;
Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 PSGNC 8
Db 9 PSGVC 13

RESULT 5
MARI ALTSP STANDARD; PRT; 14 AA.
ID MARI ALTSP
AC P29399;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Marinosatin C-2 (Marinosatin C-1; Marinosatin D).
OS Alteromonas sp. (strain B-10-31).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Alteromonas.
OX NCBI_TaxID=29456;
RN [1]
RP SEQUENCE, AND ACTIVE SITE.
RX MEDLINE=92176155; PubMed=1794974;
RA Takano R., Imada C., Kamei K., Hara S.;
RT "The reactive site of marinosatin, a proteinase inhibitor from
marine Alteromonas sp. B-10-31.";
RL J. Biochem. 110:856-858(1991).
CC -!- FUNCTION: INHIBITS SUBSTITISIN, CHYMOTRYPSIN, AND ELASTASE, BUT
NOT TRYPSIN.
KW Serine protease inhibitor; Pyrrolidone carboxylic acid.
FT PEPTIDE 1 14 MARINOSTATIN C-2.
FT PEPTIDE 3 14 MARINOSTATIN C-1.
FT PEPTIDE 4 14 MARINOSTATIN D.
FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT ACT_SITE 6 7 REACTIVE BOND.
FT SEQUENCE 14 AA; 1644 MW; 6E7CEEF92EF32E44 CRC64;

Query Match 26.1%; Score 23; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RYPS 5
Db 7 RYPS 10

RESULT 6
UC14 MAIZE STANDARD; PRT; 15 AA.
ID UC14 MAIZE
AC P80620;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 258)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED P: OF THIS UNKNOWN

CC PROTEIN IS: 5.9, ITS MW IS: 34.6 kDa.
DR Maize-2DPAGE; P80620; COLJEOPTE.
DR MaizeDB; 123944; -.
FT NON_TER 1
FT NON_TER 15
SQ SEQUENCE 15 AA; 1564 MW; CFCBBAA0B7DE6558 CRC64;

Query Match 26.1%; Score 23; DB 1; Length 15;
Best Local Similarity 57.1%; Pred. No. 7.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 GNCGLYY 12
DB 4 GDGCAYY 10

RESULT 7

SAP_STOVA STANDARD; PRT; 9 AA.
AC P24047;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Sperm-activating peptide (SAP).
OS Stomopneustes variolaris (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidea; Diadematacea; Phrynosomatoida; Stomechinidae;
OC Stomopneustes.
OX NCBI_TaxID=7663;
RN [1]
RP SEQUENCE, AND DISULFIDE BOND.
RC TISSUE=Egg jelly;
RX MEDLINE=92097763; PubMed=1756858;
RA Yoshino K.-I., Takao T., Shimonishi Y., Suzuki N.;
RT "Determination of the amino acid sequence of an intramolecular
RT disulfide linkage-containing sperm-activating peptide by tandem mass
RT spectrometry.";
RL FEBS Lett. 294:179-182(1991).
CC -!- FUNCTION: CAUSE STIMULATION OF SPERM RESPIRATION AND MOTILITY
CC THROUGH INTRACELLULAR ALKALINIZATION, TRANSIENT ELEVATIONS OF
CC CAMP, CGMP AND CLACIUM LEVELS IN SPERM CELLS, AND TRANSIENT
CC ACTIVATION AND SUBSEQUENT INACTIVATION OF THE MEMBRANE FORM OF
CC GUANYLATE CYCLASE.
FT DISULFID 3 8
SQ SEQUENCE 9 AA; 1010 MW; C469B3307B076EB9 CRC64;

Query Match 25.0%; Score 22; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSQNC 8
DB 4 PEQKC 8

RESULT 8

CXA2_CONGE STANDARD; PRT; 13 AA.
AC P01520;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-conotoxin GI2.
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE.
RX MEDLINE=8119-854; PubMed=7014556;
RA Gray W.R., Luque A., Olivera B.M., Barrett J., Cruz L.J.;
RT "Peptide toxins from Conus geographus venom.";

RJ J. Biol. Chem. 256:4734-4740(1981).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84280842; PubMed=6466616;
RA Gray W.R., Luque F.A., Galyean R., Atherton E., Sheppard R.C.,
RA Stone B.L., Reyes A., Alford J., McIntosh M., Olivera B.M.,
RA Cruz L.J., Rivier J.;
RT "Conotoxin GI: disulfide bridges, synthesis, and preparation of
RT iodinated derivatives.";
RJ Biochemistry 23:2796-2802(1984).
CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
CC bind to the nicotinic acetylcholine receptors (nAChR) and thus
CC inhibit them.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.

DR PIR; A01783; NTKN2G.
DR HSSP; P56973; 1B45.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation.
FT DISULFID 2 7
FT DISULFID 3 13
FT MOD RES 13 13
SQ SEQUENCE 13 AA; 1422 MW; DEEE831C39297EBD CRC64;

Query Match 25.0%; Score 22; DB 1; Length 13;
Best Local Similarity 50.0%; Pred. No. 9.1e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLYYS 13
DB 7 CGKHFS 12

RESULT 9

CXAL_CONMA STANDARD; PRT; 14 AA.
AC P01521;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-conotoxin MI (M1).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE.
RX MEDLINE=83073458; PubMed=7149738;
RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
RT "Isolation and structure of a peptide toxin from the marine snail
RT Conus magus.";
RL Arch. Biochem. Biophys. 218:329-334(1982).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin MI. Disulfide bonding and conformational states.";
RJ J. Biol. Chem. 258:12247-12251(1983).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.

DR PIR; A01784; NTKN1M.
DR HSSP; P56973; 1B45.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation.
FT DISULFID 3 8

FT DISULFID 4 14
FT MOD RES 14 14 AMIDATION.
SQ SEQUENCE 14 AA; 1499 MW; DEEE91898BF5E5BD CRC64;

Query Match 25.0%; Score 22; DB 1; Length 14;
Best Local Similarity 66.7%; Pred. No. 9.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYYS 13
DB 8 CGKNYS 13

RESULT 10
TRP4_LEUMA STANDARD; PRT; 9 AA.
ID TRP4_LEUMA
AC P81736;
DT 30-MAY-2003 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tachykinin-related peptide 4 (LemTRP 4).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC Biaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Midgut;
RX MEDLINE=97053012; PubMed=8897641;
RA Muren J.E., Naessel D.R.;
RT "Isolation of five tachykinin-related peptides from the midgut of
R: the cockroach Leucophaea maderae: existence of N-terminally extended
R: isoforms.";
RL Regul. Pept. 65:185-196(1996).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: MIDGUT.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD RES 9
SQ SEQUENCE 9 AA; 953 MW; 2403987699C8E5A7 CRC64;

Query Match 23.9%; Score 21; DB 1; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.3e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSNGCGL 10
DB 2 PSGFMG 8

RESULT 11
CIQA_RAT STANDARD; PRT; 15 AA.
ID CIQA_RAT
AC P31720;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Complement C1q subcomponent, A chain (Fragment).
GN CIQA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=93218657; PubMed=9464426;
RA Wing M.G., Seilly D.J., Bridgman D.J., Harrison R.A.;
RT "Rapid isolation and biochemical characterization of rat C1 and C1q.";
RL Mol. Immunol. 30:433-440(1993).
CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD

CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. THE C1Q SUBCOMPONENT IS
CC COMPOSED OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED
CC DIMERS OF THE A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-
CC LINKED DIMERS OF THE C CHAIN. IN ADDITION TO THE MAJOR A:B AND C:C
CC DIMER BANDS, RAT, UNLIKE HUMAN C1Q, CONTAINED MINOR DIMER SPECIES.
CC -!- SIMILARITY: Contains 1 collagenous domain.
DR InterPro; IPR001073; C1q.
KW PROSITE; PS01113; C1Q; PARTIAL.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat.
FT DISULFID 4 4 INTERCHAIN (WITH C-4 IN B CHAIN).
FT NCN TER 15 15
SQ SEQUENCE 15 AA; 1488 MW; 1B3D8000B7793965 CRC64;

Query Match 23.9%; Score 21; DB 1; Length 15;
Best Local Similarity 44.4%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RYPSNCGL 10
DB 5 RAPNGKDG 13

RESULT 12
FGF1_CANFA STANDARD; PRT; 15 AA.
ID FGF1_CANFA
AC P18651;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heparin-binding growth factor 1 (HBGF-1) (Acidic fibroblast growth
DE factor) (AFGF) (Alpha-endothelial cell growth factor) (Fragment).
GN FGF1 OR FGF-1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=89231704; PubMed=2714282;
RA Quinkler W., Maasberg M., Bernotat-Danielowski S., Luethe N.,
RA Sharma H.S., Schaper W.;
RT "Isolation of heparin-binding growth factors from bovine, porcine and
RT canine hearts.";
RL Eur. J. Biochem. 181:67-73(1989).
CC -!- FUNCTION: THE HEPARIN-BINDING GROWTH FACTORS ARE ANGIOGENIC AGENTS
CC IN VIVO AND ARE POTENT MITOGENS FOR A VARIETY OF CELL TYPES IN
CC VITRO. THERE ARE DIFFERENCES IN THE TISSUE DISTRIBUTION AND
CC CONCENTRATION OF THESE 2 GROWTH FACTORS.
CC -!- SUBUNIT: Monomer.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS HEPARIN, ALTHOUGH LESS STRONGLY
CC THAN DOES BFGF.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
DR PIR; S03955; S03955.
DR InterPro; IPR002348; IL1_HBGF.
DR PROSITE; PS00247; HBGF_FGF; PARTIAL.
KW Growth factor; Mitogen; Angiogenesis; Heparin-binding.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1732 MW; 53CC9A3CADDDEAA1 CRC64;

Query Match 23.9%; Score 21; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 LYYSSG 15
DB 8 LYXSNG 13

```
RESULT 13
PPCK_FASHE
ID PPCK_FASHE STANDARD; PRT; 10 AA.
AC P80525;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative phosphoenolpyruvate carboxylase (GTP) (EC 4.1.1.32)
DE (Phosphoenolpyruvate carboxylase) (PEPCK) (Newly excysted juvenile
DE protein 1) (Fragment).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fasciolidae; Fasciolidae; Fasciola.
OX NCBI_TaxID=6192;
RN [1]
RP SEQUENCE.
RX MEDLINE=95366993; PubMed=7639732;
RA Tkalcovic J., Ashman K., Meusen E.;
RT "Fasciola hepatica: rapid identification of newly excysted juvenile
RT proteins."
RL Biochem. Biophys. Res. Commun. 213:169-174 (1995).
CC -!- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -!- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE (GTP)
CC FAMILY.
DR InterPro; IPR000364; PEP carboxykin.
DR PROSITE; PS00505; PEPCK_GTP; PARTIAL.
KW Lyase; Decarboxylase; GTP-binding.
FT NON TER 10
FT SEQUENCE 10 AA; 1069 MW; 8393A6187AA9C87A CRC64;
SQ SEQUENCE 10 AA; 1069 MW; 8393A6187AA9C87A CRC64;

Query Match 22.7%; Score 20; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 YPSG 6
DB 1 YPDG 4

RESULT 14
TKL1_LOCM1
ID TKL1_LOCM1 STANDARD; PRT; 9 AA.
AC P16223;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Locustatachykinin I (TK-I).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RX TISSUE=Brain;
RX MEDLINE=90184489; PubMed=2311766;
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT "Locustatachykinin I and II, two novel insect neuropeptides with
RT homology to peptides of the vertebrate tachykinin family."
RL FEBS Lett. 261:397-401 (1990).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. STIMULATES THE CONTRACTION OF THE
CC OVIDUCT AND FOREGUT.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW tachykinin; Neuropeptide; Amidation.
FT MOD RES 9
FT SEQUENCE 9 AA; 939 MW; 2389C86B59C865A7 CRC64;
SQ SEQUENCE 9 AA; 939 MW; 2389C86B59C865A7 CRC64;

Query Match 21.6%; Score 19; DB 1; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.3e+05;
```

```
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 PSNGNGL 10
DB 2 PSGFYGV 8

RESULT 15
HTF_HELZE
ID HTF_HELZE STANDARD; PRT; 10 AA.
AC P16353;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypertrehalosaemic hormone (HeZ-HRTH).
DE Heliothis zea (Corn earworm) (Bollworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Helicoverpa.
OX NCBI_TaxID=7113;
RN [1]
RP SEQUENCE.
RX TISSUE=Corpora cardiaca;
RX MEDLINE=88326324; PubMed=3415690;
RA Gafie H., Raina A.K., Riley C.T., Fraser B.A., Bird T.G.,
RA Tseng C.M., Zhang Y.S., Hayes D.K.;
RT "Isolation and primary structure of a neuropeptide hormone from
RT Heliothis zea with hypertrehalosemic and adipokinetic activities."
RL Biochem. Biophys. Res. Commun. 155:344-350 (1988).
CC -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A31571; A31571.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
FT MOD RES 1
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT SEQUENCE 10 AA; 1096 MW; 8E70367865A5B9D1 CRC64;
SQ SEQUENCE 10 AA; 1096 MW; 8E70367865A5B9D1 CRC64;

Query Match 21.6%; Score 19; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 10 LYYSSG 15
DB 2 LTFSSG 7

Search completed: November 5, 2003, 16:57:50
Job time : 13 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Query protein - protein search, using sw model

Run on: November 5, 2003, 16:54:56 ; Search time 33 seconds
(without alignments)
117.297 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 256052604 residues

Total number of hits satisfying chosen parameters: 3526

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp Vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	30.7	10	10 P82438	P82438 nicotiana t
2	25	28.4	13	5 Q8T6E9	Q8T6E9 drosophila
3	25	28.4	13	5 Q8ST15	Q8ST15 drosophila
4	24	27.3	9	4 Q96778	Q96778 homo sapien
5	24	27.3	11	2 Q9A1Y6	Q9A1Y6 carsolella
6	24	27.3	11	5 Q9N165	Q9N165 ascaris suu
7	23	26.1	14	13 Q8JJ80	Q8JJ80 streptopeli
8	22	25.0	10	6 Q8WP04	Q8WP04 ateles belz
9	22	25.0	11	13 Q9CWA2	Q9CWA2 gallus gall
10	22	25.0	15	4 Q00604	Q00604 homo sapien
11	22	25.0	15	8 Q95770	Q95770 cyclura ric
12	22	25.0	15	13 Q9PR29	Q9PR29 micropogoni
13	21	23.9	10	13 Q90Y93	Q90Y93 gallus gall
14	21	23.9	13	4 Q9UPE7	Q9UPE7 homo sapien
15	21	23.9	13	11 Q8CJ33	Q8CJ33 mus musculu
16	21	23.9	15	10 Q9S8V7	Q9S8V7 triticum ae

17	20	22.7	10	4 Q8NEY9	Q8NEY9 homo sapien
18	20	22.7	12	2 Q93UU4	Q93UU4 escherichia
19	20	22.7	12	4 Q9UMQ9	Q9UMQ9 homo sapien
20	20	22.7	13	11 Q9WTZ6	Q9WTZ6 mus musculu
21	20	22.7	13	11 P82808	P82808 rattus norv
22	20	22.7	14	2 Q45876	Q45876 clostridium
23	20	22.7	14	2 Q45872	Q45872 clostridium
24	20	22.7	14	12 Q9PY99	Q9PY99 murine hepa
25	20	22.7	15	4 Q9UJ51	Q9UJ51 homo sapien
26	20	22.7	15	12 Q86869	Q86869 lymphocytic
27	20	22.7	15	12 Q86865	Q86865 lymphocytic
28	20	22.7	15	12 Q86867	Q86867 lymphocytic
29	19	21.6	7	10 Q9C5B3	Q9C5B3 arabidopsis
30	19	21.6	8	7 Q95213	Q95213 oryctolagus
31	19	21.6	10	4 Q15342	Q15342 homo sapien
32	19	21.6	10	4 Q9UCRC	Q9UCRC homo sapien
33	19	21.6	13	4 Q9UC27	Q9UC27 homo sapien
34	19	21.6	15	1 Q9JWH6	Q9JWH6 thermococcu
35	19	21.6	15	10 Q9S8L4	Q9S8L4 zea mays (m
36	18	20.5	9	6 Q9TRU7	Q9TRU7 bos taurus
37	18	20.5	9	11 Q9QWT0	Q9QWT0 mus musculu
38	18	20.5	9	11 Q08979	Q08979 mus musculu
39	18	20.5	10	2 Q9X533	Q9X533 escherichia
40	18	20.5	10	2 Q9X534	Q9X534 leclercia a
41	18	20.5	10	2 Q48469	Q48469 klebsiella
42	18	20.5	11	5 Q9TWM2	Q9TWM2 aplysia cal
43	18	20.5	11	12 Q86866	Q86866 lymphocytic
44	18	20.5	11	12 Q86864	Q86864 lymphocytic
45	18	20.5	11	12 Q86868	Q86868 lymphocytic

ALIGNMENTS

RESULT 1

P82438
ID P82438 PRELIMINARY; PRT; 10 AA.
AC P82438;
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE 50 kDa cell wall protein (Fragment).
OS Nicotiana tabacum (Common tobacco).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
CC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. PETIT HAVANA;
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA Wojtaszek P., Bolwell G.P.;
RT "Proteomic study of secondary cell wall proteins from transformed
RT tobacco culture."
RL Planta C:0-0(2000).
CC -!- SUBCELLULAR LOCATION: CELL WALL.
CC -!- TISSUE SPECIFICITY: XYLEM.
KW Cell wall.
FT NON TER 10 10
SQ SEQUENCE 10 AA; 1126 MW; C68E32486AF77B46 CRC64;
Query Match 30.7%; Score 27; DB 10; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RYPSGN 7

Db 3 QYPXGN 8

RESULT 2

Q8T6E9
ID Q8T6E9 PRELIMINARY; PRT; 13 AA.

```
AC Q8T659;
CT 01-JUN-2002 (TrEMBLrel. 21, Created)
CT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
CT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Myocyte enhancing factor 2 (Fragment).
GN MEF2.
CS Drosophila miranda (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7229;
RN [1]
RP SEQUENCE FROM N.A.
RA Schaeffer S.W., Goetting-Minesky M.P., Kovacevic M., Peoples J.,
RA Graybill J.L., Miller J.M., Kim K., Nelson J.G., Anderson W.W.;
RT "Evolutionary genomics of inversions in Drosophila pseudoobscura:
RT Modes of selection.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF476817; AAL91815.1; -
DR FlyBase; FBgn0062329; Dmir\Mef2.
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1313 MW; 15BF380B6BE05050 CRC64;

Query Match 29.4%; Score 25; DB 5; Length 13;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPSGN 7
Db 8 YPSGS 12

RESULT 3
Q8STI5 PRELIMINARY; PRT; 13 AA.
AC Q8STI5;
CT 01-JUN-2002 (TrEMBLrel. 21, Created)
CT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
CT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Myocyte enhancing factor 2 (Fragment).
GN MEF2.
CS Drosophila pseudoobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7237;
RN [1]
RP SEQUENCE FROM N.A.
RA Schaeffer S.W., Goetting-Minesky M.P., Kovacevic M., Peoples J.,
RA Graybill J.L., Miller J.M., Kim K., Nelson J.G., Anderson W.W.;
RT "Evolutionary genomics of inversions in Drosophila pseudoobscura:
RT Modes of selection.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF476728; AAL91726.1; -
DR EMBL; AF476729; AAL91727.1; -
DR EMBL; AF476730; AAL91728.1; -
DR EMBL; AF476731; AAL91729.1; -
DR EMBL; AF476732; AAL91730.1; -
DR EMBL; AF476733; AAL91731.1; -
DR EMBL; AF476734; AAL91732.1; -
DR EMBL; AF476735; AAL91733.1; -
DR EMBL; AF476736; AAL91734.1; -
DR EMBL; AF476737; AAL91735.1; -
DR EMBL; AF476738; AAL91736.1; -
DR EMBL; AF476739; AAL91737.1; -
DR EMBL; AF476740; AAL91738.1; -
DR EMBL; AF476741; AAL91739.1; -
DR EMBL; AF476742; AAL91740.1; -
DR EMBL; AF476743; AAL91741.1; -
DR EMBL; AF476744; AAL91742.1; -
DR EMBL; AF476745; AAL91743.1; -
DR EMBL; AF476746; AAL91744.1; -
DR EMBL; AF476747; AAL91745.1; -
DR EMBL; AF476748; AAL91746.1; -
DR EMBL; AF476749; AAL91747.1; -
DR EMBL; AF476750; AAL91748.1; -
DR EMBL; AF476751; AAL91749.1; -
DR EMBL; AF476752; AAL91750.1; -
DR EMBL; AF476753; AAL91751.1; -
DR EMBL; AF476754; AAL91752.1; -
DR EMBL; AF476755; AAL91753.1; -
DR EMBL; AF476756; AAL91754.1; -
DR EMBL; AF476757; AAL91755.1; -
DR EMBL; AF476758; AAL91756.1; -
DR EMBL; AF476759; AAL91757.1; -
DR EMBL; AF476760; AAL91758.1; -
DR EMBL; AF476761; AAL91759.1; -
DR EMBL; AF476762; AAL91760.1; -
DR EMBL; AF476763; AAL91761.1; -
DR EMBL; AF476764; AAL91762.1; -
DR EMBL; AF476765; AAL91763.1; -
DR EMBL; AF476766; AAL91764.1; -
DR EMBL; AF476767; AAL91765.1; -
DR EMBL; AF476768; AAL91766.1; -
DR EMBL; AF476769; AAL91767.1; -
DR EMBL; AF476770; AAL91768.1; -
DR EMBL; AF476771; AAL91769.1; -
DR EMBL; AF476772; AAL91770.1; -
DR EMBL; AF476773; AAL91771.1; -
DR EMBL; AF476774; AAL91772.1; -
DR EMBL; AF476775; AAL91773.1; -
DR EMBL; AF476776; AAL91774.1; -
DR EMBL; AF476777; AAL91775.1; -
DR EMBL; AF476778; AAL91776.1; -
DR EMBL; AF476779; AAL91777.1; -
DR EMBL; AF476780; AAL91778.1; -
DR EMBL; AF476781; AAL91779.1; -
DR EMBL; AF476782; AAL91780.1; -
DR EMBL; AF476783; AAL91781.1; -
DR EMBL; AF476784; AAL91782.1; -
DR EMBL; AF476785; AAL91783.1; -
DR EMBL; AF476786; AAL91784.1; -
DR EMBL; AF476787; AAL91785.1; -
DR EMBL; AF476788; AAL91786.1; -
DR EMBL; AF476789; AAL91787.1; -
DR EMBL; AF476790; AAL91788.1; -
DR EMBL; AF476791; AAL91789.1; -
DR EMBL; AF476792; AAL91790.1; -
DR EMBL; AF476793; AAL91791.1; -
DR EMBL; AF476794; AAL91792.1; -
DR EMBL; AF476795; AAL91793.1; -
DR EMBL; AF476796; AAL91794.1; -
DR EMBL; AF476797; AAL91795.1; -
DR EMBL; AF476798; AAL91796.1; -
DR EMBL; AF476799; AAL91797.1; -
DR EMBL; AF476800; AAL91798.1; -
DR EMBL; AF476801; AAL91799.1; -
DR EMBL; AF476802; AAL91800.1; -
DR EMBL; AF476803; AAL91801.1; -
DR EMBL; AF476804; AAL91802.1; -
DR EMBL; AF476805; AAL91803.1; -
DR EMBL; AF476806; AAL91804.1; -
DR EMBL; AF476807; AAL91805.1; -
DR EMBL; AF476808; AAL91806.1; -
DR EMBL; AF476809; AAL91807.1; -
DR EMBL; AF476810; AAL91808.1; -
DR EMBL; AF476811; AAL91809.1; -
DR EMBL; AF476812; AAL91810.1; -
DR EMBL; AF476813; AAL91811.1; -
DR EMBL; AF476814; AAL91812.1; -
DR EMBL; AF476815; AAL91813.1; -
DR EMBL; AF476816; AAL91814.1; -
DR FlyBase; FBgn0062298; Opse\Mef2.
FT NON_TER 1
FT NON_TER 13
```

SQ SEQUENCE 13 AA; 1313 MW; 15BF380B6BE05050 CRC64;
Query Match: 28.4%; Score 25; DB 5; Length 13;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 YPSGN 7
DQ 8 YPSGS 12
RESULT 4
Q96T78 PRELIMINARY; PRT; 9 AA.
AC Q96T78;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Contactin-associated protein 2 (Fragment).
GN CNTNAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. PubMed=11352571;
RX MEDLINE=21250995; Scherer S.W.;
RA Nakabayashi K.;
RT "The human contactin-associated protein 2 (CNTNAP2) spans over 2 Mb of
RT DNA at chromosome 7q35.";
RL Genomics 73:108-112(2001).
DR EMBL; AF318295; AAK49906.1; -.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 1080 MW; 9139A2D5A77B51EA CRC64;

Query Match: 27.3%; Score 24; DB 4; Length 9;
Best Local Similarity 80.0%; Pred. No. 8.3e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRYPS 5
DQ 2 DRYPS 6

RESULT 5
Q9AIY6 PRELIMINARY; PRT; 11 AA.
AC Q9AIY6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Tryptophanyl-tRNA synthetase (Fragment).
GN TRPS.
OS Carsonella ruddii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Candidatus Carsonella.
OX NCBI_TaxID=114186;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20336438; PubMed=10877784;
RA Thao M.L., Moran N.A., Abbot P., Brennan E.B., Burkhardt D.H.,
RA Baumann P.;
RT "Cospeciation of psyllids and their primary prokaryotic
RT endosymbionts.";
RL Appl. Environ. Microbiol. 66:2898-2905(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21125546; PubMed=11222582;
RA Clark M.A., Baumann L., Thao M.L., Moran N.A., Baumann P.;
RT "Degenerative Minimalism in the Genome of a Psyllid Endosymbiont.";
RL J. Bacteriol. 183:1853-1861(2001).
DR EMBL; AF211138; AAK15388.1; -.
KW Aminoacyl-tRNA synthetase.

FT NON_TER 1
SQ SEQUENCE 11 AA; 1295 MW; 0CA993A5345B5720 CRC64;
Query Match: 27.3%; Score 24; DB 2; Length 11;
Best Local Similarity 42.9%; Pred. No. 2e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 7 NCGLYS 13
DQ 1 NCSTLYN 7

RESULT 6
Q9NL65 PRELIMINARY; PRT; 11 AA.
AC Q9NL65;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ASABF-delta (Fragment).
GN ASABF-DELTA.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato Y.;
RT "Ascaris suum asabf-delta gene, exon 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB029815; BAA89496.1; -.
FT NON_TER 1
FT NON_TER 11
SQ SEQUENCE 11 AA; 1187 MW; 8EADDDCCDEAB5861 CRC64;

Query Match: 27.3%; Score 24; DB 5; Length 11;
Best Local Similarity 56.7%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 NCGLY 12
DQ 1 NCSTGY 6

RESULT 7
Q8J80 PRELIMINARY; PRT; 14 AA.
AC Q8J80;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Beta fibrinogen (Fragment).
OS Streptopelia bitorquata.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae;
OC Streptopelia.
OX NCBI_TaxID=177146;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson K.P., de Kort S., Dinwoodey K., Mateman A.C., ten Cate C.,
RA Lessells C.M., Clayton D.H.;
RT "A molecular phylogeny of the dove genera Streptopelia and Columba.";
RL Auk 118:874-887(2001).
DR EMBL; AF353457; AAM22348.1; -.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
FT NON_TER 1
FT NON_TER 14
SQ SEQUENCE 14 AA; 1655 MW; 9AF33468DACD4D41 CRC64;

Query Match: 26.1%; Score 23; DB 13; Length 14;
Best Local Similarity 50.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 9 GLYYS 14
|:|:|:
Db 2 GMYFST 7

RESULT 8
Q8WFO4 Q8WPC4 PRELIMINARY; PRT; 10 AA.

AC Q8WPC4;
CT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Oculocutaneous albinism type II (Fragment).
GN OCA2.
OS Ateles belzebuth chamek (Chamek spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxID=118643;
RN {}
RP SEQUENCE FROM N.A.
RX MEDLINE=21636947; PubMed=11778586;
RA Seanez H.N., Lima C.R., Lemos B., Bonvicino C.R., Moreira M.A.X.,
RA Canavez F.C.;
RT "Gene assignment in Ateles paniscus chamek (Platyrrhini, Primates).
RT Allocation of 18 markers of human syntenic groups 1,2,7,14,15,17 and
RT 22.";
RL Chromosome Res. 9:631-639(2001).
DR EMBL; AF375652; AAC33489.1; -.
FT NON_TER 1
SQ SEQUENCE 10 AA; 901 MW; 22DF477DD87EA5B8 CRC64;

Query Match 25.0%; Score 22; DB 6; Length 10;
Best Local Similarity 75.0%; Pred. No. 4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 GNCG 9
|:|:
Db 3 GSCG 6

RESULT 9
Q90WA2 Q90WA2 PRELIMINARY; PRT; 11 AA.

AC Q90WA2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Extracellular fatty acid binding protein (Fragment).
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN {}
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Wang Q., Li N., Deng X., Li H.;
RT "Single Nucleotide Polymorphism Analysis on Encoding Region of
RT Extracellular Fatty Acid Binding Protein Genes and Their Associations
RT With the Fattiness Trait in Chicken."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF402001; AAK94062.1; -.
FT NON_TER 1
FT NON_TER 11
SQ SEQUENCE 11 AA; 1277 MW; 873417F02B187AAA CRC64;

Query Match 25.0%; Score 22; DB 13; Length 11;
Best Local Similarity 100.0%; Pred. No. 4.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LYYS 13
|:|:
Db 2 NCRLHSTS 9

Db 9 LYYS 11

RESULT 10
Q0C604 Q0C604 PRELIMINARY; PRT; 15 AA.

AC Q0C604;
ET 01-JUL-1997 (TREMBLrel. 04, Created)
ET 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Type I collagen alpha 1(I) chain (Fragment).
GN COL1A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN {}
RP SEQUENCE FROM N.A.
RX MEDLINE=96336021; PubMed=8757037;
RA Mayer S.A., Rubin B.S., Starman B.J., Byers P.H.;
RT "Spontaneous multivessel cervical artery dissection in a patient with
RT a substitution of alanine for glycine (G13A) in the alpha 1 (I) chain
RT of type I collagen.";
RL Neurology 47:552-556(1996).
DR EMBL; S83315; AAB50780.2; -.
KW Collagen.
FT NON_TER 1
FT NON_TER 15
SQ SEQUENCE 15 AA; 1413 MW; 4CA8D720AAF780FC CRC64;

Query Match 25.0%; Score 22; DB 4; Length 15;
Best Local Similarity 71.4%; Pred. No. 6e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSGNCG 10
|:|:|:
Db 2 PSGPRGL 8

RESULT 11
Q95770 Q95770 PRELIMINARY; PRT; 15 AA.

AC Q95770;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
GN ND4.
OS Cyclura ricordi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Iguanidae; Iguaninae; Cyclura.
OX NCBI_TaxID=51215;
RN {}
RP SEQUENCE FROM N.A.
RX MEDLINE=97019047; PubMed=8865663;
RA Sites J.W. Jr., Davis S.K., Guerra T., Iverson C.B., Shell H.L.;
RT "Character congruence and phylogenetic signal in molecular and
RT morphological data sets: a case study in the living Iguanas (Squamata,
RT Iguanidae).";
RL Mol. Biol. Evol. 13:1087-1105(1996).
DR EMBL; U66237; AAB07473.1; -.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 15 AA; 1715 MW; 8327178E7927A57E CRC64;

Query Match 25.0%; Score 22; DB 8; Length 15;
Best Local Similarity 50.0%; Pred. No. 6e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 NCGLYSS 14
|:|:|:
Db 2 NCRLHSTS 9

RESULT 12

C9PRZ9
ID Q9PRZ9 PRELIMINARY; PRT; 15 AA.
AC Q9PRZ9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Gonadotropin I beta subunit (Fragment).
CS Micropogonias undulatus (Atlantic croaker).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
CC Sciaenidae; Micropogonias.
CX NCBI_TaxID=29154;
RN [1]
RP SEQUENCE.
RX MEDLINE=94010173; PubMed=8405893;
RA Copeiland P.A., Thomas P.;
RT "Isolation of gonadotropin subunits and evidence for two distinct
gonadotropins in Atlantic croaker (Micropogonias undulatus).";
EL Gen. Comp. Endocrinol. 91:115-125(1993).
SQ SEQUENCE 15 AA; 1732 MW; 81709992C3D86A4A CRC64;

Query Match 25.0%; Score 22; DB 13; Length 15;
Best Local Similarity 75.0%; Pred. No. 6e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLY 11
Db 4 CGFY 7

RESULT 13

Q90Y93
ID Q90Y93 PRELIMINARY; PRT; 10 AA.
AC Q90Y93;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Growth hormone (Fragment).
GN GH.
OS Gallus gallus (Chicken).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC Gallus.
CX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Kansaku N., Nakada A., Yagi E., Okabayashi H., Guemene D.;
RT "Genetic variation of chicken growth hormone gene.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB061722; BAB69037.1; -.
FT NON_TER 1
FT NON_TER 10
SQ SEQUENCE 10 AA; 1155 MW; 6841751775A40AAB CRC64;

Query Match 23.9%; Score 21; DB 13; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSG 6
Db 4 DRSRPG 9

RESULT 14

Q9UPE7
ID Q9UPE7 PRELIMINARY; PRT; 13 AA.
AC Q9UPE7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Inosine monophosphatase 2 (Fragment).
GN IMPA2.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97463449; PubMed=93222233;
RA Yoshikawa T., Turner G., Esterling L.E., Sanders A.R.,
RA Detera-Wadleigh S.D.;
RT "A novel human myo-inositol monophosphatase gene, IMP18p, maps to a
susceptibility region for bipolar disorder.";
RL Mol. Psychiatry 2:393-397(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20284187;
RA Yoshikawa T., Padigaru M., Karkera J.D., Sharma M., Berrettini W.H.,
RA Esterling L.E., Detera-Wadleigh S.D.;
RT "Genomic structure and novel variants of myo-inositol monophosphatase
2.";
RL Mol. Psychiatry 5:165-171(2000).
DR EMBL; AF025882; AAD22136.1; -.
DR EMBL; AF025881; AAD22136.1; JOINED.
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1589 MW; F3415D841F48D401 CRC64;

Query Match 23.9%; Score 21; DB 4; Length 13;
Best Local Similarity 60.0%; Pred. No. 7.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPS 5
Db 1 ERFPS 5

RESULT 15

Q8CU33
ID Q8CU33 PRELIMINARY; PRT; 13 AA.
AC Q8CU33;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Myoneurin (Fragment).
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Kidney;
RX MEDLINE=20334280; PubMed=10873615;
RA Alliel P.M., Seddigi N., Goudou D., Cifuentes-Diaz C., Romero N.,
RA Velasco E., Rieger F., Perin J.P.;
RT "Myoneurin, a novel member of the BTB/POZ-zinc finger family highly
expressed in human muscle.";
RL Biochem. Biophys. Res. Commun. 273:385-391(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Kidney;
RA Bitoun M., Perin J.P., Seddigi N., Goudou D., Camuzat A., Matter M.G.,
RA Rieger F., Alliel P.M.;
RT "The human and mouse myoneurin genes: Genomic organization, splice
variants, chromosomal mapping and flanking genes.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF530454; AAN28707.1; -.
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1445 MW; 9CB68B708CA1587E CRC64;

Query Match 23.9%; Score 21; DB 11; Length 13;

Best Local Similarity 57.1%; Pred. No. 7.6e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 YPSGNCG 9
| | | |
Db 4 YICGICG 10

Search completed: November 5, 2003, 16:58:37
Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 16:48:21 ; Search time 41 Seconds
(without alignments)
58.071 Million cell; updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRYPGKGLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 350435

Minimum DB seq length: 3

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq 19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	15	20	AAW88237 Human prothrombina
2	37	42.0	14	23	AAE28619 Human K-betaV2 pro
3	33	37.5	9	22	AAU23853 Human MHC class I
4	33	37.5	9	22	AAU24319 Human MHC class I
5	33	37.5	9	22	AAU24423 Human MHC molecule
6	33	37.5	9	22	AAU24441 Human MHC molecule
7	33	37.5	10	22	AAU24484 Human MHC molecule
8	33	37.5	12	21	AAU26044 Human IgE C-epsilo
9	33	37.5	12	22	AAU16769 Peptide EED16/47/4

10	33	37.5	12	23	ABJ00390 Human IgE cyclic i
11	33	37.5	15	18	AAW23509 Purified cis-9,10-
12	33	37.5	15	19	AAW57788 Fatty acid amide h
13	31	35.2	10	22	AAG96214 Human complementar
14	31	35.2	15	5	AAP40328 Sequence of conoto
15	31	35.2	15	16	AAW75265 Alpha-conotoxin GI
16	31	35.2	15	18	AAW12727 Alpha-conotoxin pe
17	30.5	34.7	15	24	ABR34060 Human cancer-relat
18	30.5	34.7	15	24	ABR34207 Human cancer-relat
19	30	34.1	7	20	AAV30230 Physarum polyceph
20	30	34.1	10	22	AAG94296 Human complementar
21	30	34.1	12	20	AAV29882 Humicola lanuginos
22	30	34.1	13	15	AAAR46829 pH 2.5 acid phosph
23	30	34.1	14	15	AAAR58323 Hypotensive polype
24	30	34.1	14	19	AAW59655 Amino acid sequenc
25	30	34.1	14	22	AAW68153 Peptide derived fr
26	30	34.1	14	22	AAW68154 Peptide derived fr
27	30	34.1	14	22	AAW68187 Peptide derived fr
28	30	34.1	14	22	AAW68188 Peptide derived fr
29	30	34.1	15	16	AAW63873 Cryoglobulinemia t
30	30	34.1	15	19	AAW54431 Human PS12 protei
31	30	34.1	15	20	AAV30227 Physarum polyceph
32	30	34.1	15	21	AAW8421 Antigen peptide de
33	30	34.1	15	22	AAW78001 Human actin 14 pep
34	29	33.0	10	22	AAU24373 Human MHC class I
35	29	33.0	10	22	AAU24478 Human MHC molecule
36	29	33.0	14	14	AAW69320 Gp IIb/IIIa recept
37	29	33.0	14	21	AAW11799 Human ERV2 protein
38	29	33.0	14	21	AAW95520 Fibrin polymerizat
39	29	33.0	15	20	AAW07427 HSV-1 TK amino aci
40	29	33.0	15	20	AAW97159 HSV-1 TK amino aci
41	29	33.0	15	21	AAW93038 Transforming growt
42	29	33.0	15	24	ABG99218 Thymidine kinase (
43	28	31.8	6	24	ABP56238 Targeting moiety p
44	28	31.8	7	23	AAU11469 Peptide for orthog
45	28	31.8	8	19	AAW63298 Peptide which is n

ALIGNMENTS

RESULT 1

AAW88237

ID AAW88237 standard; Peptide: 15 AA.

XX

AC AAW88237;

XX

DT 15-MAR-1999 (first entry)

XX

DE Human prothrombinase Fgl2 epitope.

XX

KW Prothrombinase; hfg12; human; immune coagulation; antibody;

KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;

KW gastrointestinal disease; foetal loss; therapy; vaccine; epitope.

XX

CS Homo sapiens.

XX

PN WO9851335-A1.

XX

PD 19-NOV-1998.

XX

PF 15-MAY-1998; 98WO-CA00475.

XX

PR 10-OCT-1997; 97US-0061684.

PR 15-MAY-1997; 97US-0046537.

PA (LEVY/) LEVY G.

XX

PI Levy G;

XX

DR WFI: 1999-059687/05.

XX

PT Modulating immune coagulation - by using Fgl2 antibodies and

PT compounds, used to treat conditions including graft rejection and
PT foetal loss
XX
PS Claim 4; Page 72; 105pp; English.
XX
CC This peptide corresponds to amino acid residues 364-378 of human
CC prothrombinase Fgl2 (see AAW88235). A claimed method of preventing
CC or treating a condition requiring a reduction in immune
CC coagulation comprises administering an inhibitor of Fgl2. The
CC inhibitor is preferably an antibody that binds to the Fgl2 epitope.
CC The condition to be treated in graft rejection of foetal loss
CC (claimed).
XX
SQ Sequence 15 AA;
Query Match 100.0%; Score 88; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRYPSGNCGLYSSG 15
Db 1 DRYPSGNCGLYSSG 15
RESULT 2
AAE28619
ID AAE28619 standard; peptide; 14 AA.
XX
AC AAE28619;
XX
DT 27-DEC-2002 (first entry)
XX
DE Human K+betaM2 protein asparagine glycosylation site #1.
XX
KW Human: potassium channel beta-subunit; K+betaM2 protein; neural disorder;
KW reproductive disorder; metabolic disorder; premature puberty; nephritis;
KW endocrine disorder; memory disorder; neuroendocrine condition; asthma;
KW spermatogenesis; renal disease; learning deficiency; Alzheimer's disease;
KW neurodegenerative disease; proliferative disorder; autoimmune disease;
KW carcinoma tumour; blood coagulation disease; blood platelet disease;
KW rheumatoid arthritis; allergy; hyperproliferative disease; gene therapy;
KW graft-versus-host disease; organ rejection; antisterility; thrombolytic;
KW antiinflammatory; neuroprotective; anti-Parkinsonian; immunosuppressive;
KW nephrotropic; cytostatic; nootropic; hypotensive; vulnerary.
XX
OS Homo sapiens.
XX
XX WO200266601-A2.
XX
XX 29-AUG-2002.
XX
XX 24-JAN-2002; 2002WO-US02332.
XX
XX 24-JAN-2001; 2001US-263872P.
XX
XX 14-FEB-2001; 2001US-269794P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Feder J, Lee L, Chen J, Jackson D, Ramanathan C, Siemers N;
PI Chang H, Carroll P;
XX
XX WPI; 2002-691617/74.
XX
XX New potassium channel beta-subunit, K+betaM2, proteins and nucleic
PT acids, useful for diagnosing, treating and/or preventing e.g.
PT reproductive, neural, metabolic, endocrine, memory, neurodegenerative
PT disorders or diseases -
XX
PS Disclosure; Page 354; 366pp; English.
XX
CC The present invention relates to human potassium channel: beta-subunit
CC (K+betaM2) proteins and polynucleotides encoding such proteins. The
CC K+betaM2 sequences are useful for diagnosing, treating and/or preventing

CC reproductive disorders, neural disorders, disorders related to aberrant
CC potassium regulation or hyper potassium channel activity, metabolic
CC disorders (e.g. premature puberty), endocrine disorders (e.g. aberrant
CC growth hormone synthesis and/or secretion), memory disorder, disorders
CC of the testis (e.g. spermatogenesis), neuroendocrine condition related
CC to aberrant thyroid hormone release, renal disease or disorders (e.g.
CC nephritis), disorders related to aberrant higher brain function (e.g.
CC learning deficiencies), neurodegenerative diseases (e.g. Alzheimer's
CC disease), proliferative disorders (e.g. carcinoma tumour) and disorders
CC involving excessive smooth muscle tone or excitability (e.g. asthma).
CC They may be used to modulate haemostatic or thrombolytic activity, to
CC treat or prevent blood coagulation diseases or disorders, blood platelet
CC diseases, wounds, autoimmune diseases, disorders or conditions (e.g.
CC rheumatoid arthritis), allergic reactions (e.g. asthma), organ rejection
CC or graft-versus-host disease, and hyperproliferative diseases. K+betaM2
CC sequences are also used in gene therapy. The present sequence is human
CC K+betaM2 protein asparagine glycosylation site.
XX
SQ Sequence 14 AA;
Query Match 42.0%; Score 37; DB 23; Length 14;
Best Local Similarity 75.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 SGNCGLYY 12
Db 4 SGNCSRY 11
RESULT 3
AAU23853
ID AAU23853 standard; Peptide; 9 AA.
XX
AC AAU23853;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human XHC class I molecule HLA-A1 binding 103P2D6 peptide #38.
XX
KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
KW single chain monoclonal antibody; cervix; human.
XX
OS Homo sapiens.
XX
XX WO200162925-A2.
XX
XX 30-AUG-2001.
XX
XX 26-FEB-2001; 2001WO-US05996.
XX
XX 24-FEB-2000; 2000US-0194558.
XX
XX 13-JUL-2000; 2000US-0218856.
XX
XX (UROG-) UROGENESYS INC.
XX
XX Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
PI Chalitta-eid PM, Faris M, Jakobovits A;
XX
XX WPI; 2001-557705/62.
XX
XX New polynucleotide for treating and diagnosing prostate cancer is the
PT 103P2D6 gene which encodes for 103P2D6-related proteins -
XX
XX Example 15; Page 81; 132pp; English.
XX
CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and
CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
CC adult tissue but is aberrantly expressed in some foetal tissues and many
CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
CC related protein and also peptide fragments of the protein are therefore

CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P2D6. The sequences can be used
CC in diagnostic methods to monitor the level of 103P2D6 gene products in
CC serum, blood, urine and tissue and to thereby detect the presence of
CC cancerous cells.
XX
SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;
Best Local Similarity 75.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12
 : : ||| |
Db 2 SGRCGLGY 9

RESULT 4
AAU24319
ID AAU24319 standard; Peptide; 9 AA.
XX
AC AAU24319;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human MHC class I molecule HLA-B7 binding 103P2D6 peptide #4.
XX
KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
KW single chain monoclonal antibody; cervix; human.
XX
CS Homo sapiens.
XX
PN WC200162925-A2.
XX
PD 30-AUG-2001.
XX
PF 26-FEB-2001; 2001WO-US05996.
XX
PR 24-FEB-2000; 2000US-0184558.
PR 13-JUL-2000; 2000US-0218856.
XX
PA (UROC-) UROGENESYS INC.
XX
PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
PI Challita-eid PM, Faris M, Jakobovits A;
XX
DR WPI; 2001-557705/62.
XX
PT New polynucleotide for treating and diagnosing prostate cancer is the
PT 103P2D6 gene which encodes for 103P2D6-related proteins -
XX
PS Example 15; Page 94; 132pp; English.
XX
CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and
CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
CC adult tissue but is aberrantly expressed in some foetal tissues and many
CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
CC related protein and also peptide fragments of the protein are therefore
CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P2D6. The sequences can be used
CC in diagnostic methods to monitor the level of 103P2D6 gene products in
CC serum, blood, urine and tissue and to thereby detect the presence of

CC cancerous cells.
XX
SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;
Best Local Similarity 75.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12
 : : ||| |
Db 1 SGRCGLGY 8

RESULT 5
AAU24423
ID AAU24423 standard; Peptide; 9 AA.
XX
AC AAU24423;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human MHC molecule H2A-B35 binding 103P2D6 peptide #8.
XX
KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
KW single chain monoclonal antibody; cervix; human.
XX
CS Homo sapiens.
XX
PN WC200162925-A2.
XX
PD 30-AUG-2001.
XX
PF 26-FEB-2001; 2001WO-US05996.
XX
PR 24-FEB-2000; 2000US-0184558.
PR 13-JUL-2000; 2000US-0218856.
XX
PA (UROC-) UROGENESYS INC.
XX
PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
PI Challita-eid PM, Faris M, Jakobovits A;
XX
DR WPI; 2001-557705/62.
XX
PT New polynucleotide for treating and diagnosing prostate cancer is the
PT 103P2D6 gene which encodes for 103P2D6-related proteins -
XX
PS Example 15; Page 97; 132pp; English.
XX
CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and
CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
CC adult tissue but is aberrantly expressed in some foetal tissues and many
CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
CC related protein and also peptide fragments of the protein are therefore
CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P2D6. The sequences can be used
CC in diagnostic methods to monitor the level of 103P2D6 gene products in
CC serum, blood, urine and tissue and to thereby detect the presence of
CC cancerous cells.
XX
SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;
Best Local Similarity 75.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12
|||
Db 2 SGRGGLGY 9

RESULT 6
AAU24441

ID AAU24441 standard; Peptide; 9 AA.
XX
AC AAU24441;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human MHC molecule HLA-B35 binding 103P2D6 peptide #26.
XX
KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
KW single chain monoclonal antibody; cervix; human.
XX
OS Homo sapiens.
XX
FN WO200162925-A2.
XX
PD 30-AUG-2001.
XX
PF 26-FEB-2001; 2001WO-US05996.
XX
PR 24-FEB-2000; 2000US-0184558.
PR 13-JUL-2000; 2000US-0218856.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
PI Challita-eid PM, Faris M, Jakobovits A;
XX
DR WPI; 2001-557705/62.
XX
PT New polynucleotide for treating and diagnosing prostate cancer is the
PT 103P2D6 gene which encodes for 103P2D6-related proteins -
XX
PS Example 15; Page 98; 132pp; English.
XX
CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and
CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
CC adult tissue but is aberrantly expressed in some foetal tissues and many
CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
CC related protein and also peptide fragments of the protein are therefore
CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P2D6. The sequences can be used
CC in diagnostic methods to monitor the level of 103P2D6 gene products in
CC serum, blood, urine and tissue and to thereby detect the presence of
CC cancerous cells.
XX
SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;
Best Local Similarity 75.0%; Pred. No. 9.3e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12
|||
Db 1 SGRGGLGY 8

RESULT 7
AAU24484

ID AAU24484 standard; Peptide; 10 AA.

XX AAU24484;
AC
XX
DT 17-DEC-2001 (first entry)
XX
DE Human MHC molecule HLA-A35 binding 103P2D6 peptide #19.
XX
KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;
KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;
KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;
KW single chain monoclonal antibody; cervix; human.
XX
OS Homo sapiens.
XX
FN WO200162925-A2.
XX
PD 30-AUG-2001.
XX
PF 26-FEB-2001; 2001WO-US05996.
XX
PR 24-FEB-2000; 2000US-0184558.
PR 13-JUL-2000; 2000US-0218856.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;
PI Challita-eid PM, Faris M, Jakobovits A;
XX
DR WPI; 2001-557705/62.
XX
PT New polynucleotide for treating and diagnosing prostate cancer is the
PT 103P2D6 gene which encodes for 103P2D6-related proteins -
XX
PS Example 15; Page 99; 132pp; English.
XX
CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and
CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal
CC adult tissue but is aberrantly expressed in some foetal tissues and many
CC cancers including tumours of the prostate, testis, bladder, bone, cervix,
CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its
CC related protein and also peptide fragments of the protein are therefore
CC useful for diagnosing and treating cancer. A vector comprising a
CC polynucleotide which encodes a single chain monoclonal antibody, that
CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme
CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,
CC are both useful in the preparation of a composition for treating a
CC patient with a cancer that expresses 103P2D6. The sequences can be used
CC in diagnostic methods to monitor the level of 103P2D6 gene products in
CC serum, blood, urine and tissue and to thereby detect the presence of
CC cancerous cells.
XX
SQ Sequence 10 AA;

Query Match 37.5%; Score 33; DB 22; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12
|||
Db 2 SGRGGLGY 9

RESULT 8
AAB26044

ID AAB26044 standard; Peptide; 12 AA.
XX
AC AAB26044;
XX
DT 05-JAN-2001 (first entry)
XX
DE Human IgE C-epsilon-2 PTmAb0011 peptide ligand SEQ ID NO:138.
XX
KW Epitope; mimotope; human; immunoglobulin E; IgE; C-epsilon-2 domain;

KW allergic disease; immunophylaxis; immunotherapy; antiallergic;
KW immunosuppressive; vaccine; histamine release inhibitor; immunogen;
KW allergy; atopy.

XX Homo sapiens.
XX
PN WO200050460-A1.
XX
PD 31-AUG-2000.

XX 22-FEB-2000; 2000WO-EP01455.

XX 25-FEB-1999; 99GB-0004405.
PR 29-MAR-1999; 99GB-0007151.
PR 07-MAY-1999; 99GB-0010537.
PR 07-MAY-1999; 99GB-0010538.
PR 07-AUG-1999; 99GB-0018594.
PR 07-AUG-1999; 99GB-0018603.
PR 07-SEP-1999; 99GB-0021046.
PR 07-SEP-1999; 99GB-0021047.
PR 29-OCT-1999; 99GB-0025619.
PR 23-NOV-1999; 99GB-0027695.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA (PEPT-) PEPTIDE THERAPEUTICS LTD.

XX
PI Dyson M, Friede M, Greenwood J, Hewitt E, Lamont A, Mason S;
PI Randall R, Turnell WG, Van Vechelen MP, Virais De Bassois YC;
XX
DR WPI; 2000-572073/53.

XX Peptides useful for treating, preventing and ameliorating allergic
PT diseases, comprising an isolated surface exposed group of a specific
PT domain from immunoglobulin E -

XX Example 7; Page 53; 129pp; English.

XX The present invention describes a peptide (I) comprising an isolated
CC surface exposed group/epitope (E1) of C-epsilon-2 domain (D) of
CC immunoglobulin E (IgE), or its mimotope. Also described are: (1) an
CC for treating allergies comprising (ii); (3) a ligand (iv) capable of
CC recognising E; (4) a pharmaceutical composition (PC) comprising (iv);
CC (5) a peptide (Ia) capable of being recognised by (iv); (6) an immunogen
CC (Iia) comprising (Ia); and (7) producing (iii) by producing (ii). (i)
CC can have antiallergic and immunosuppressive activities, and can be used
CC as a vaccine and histamine release inhibitor. (i), (ii) and (iii) are
CC useful in medicine and in the manufacture of medicaments for treating
CC and preventing allergies. (iv) is useful for identifying mimotopes of Pi,
CC in medicine and also in manufacturing medicaments for treating
CC allergies. (i) is useful in diagnostics and in the affinity purification
CC of circulating anti-IgE antibodies from blood. (i), (iii) and PC are
CC useful for treating a patient susceptible to or suffering from allergies.
CC (iv) is also useful in diagnosing atopy. AAB25907 to AAB26099 represent
CC peptide sequences which are used in the exemplification of the present
CC invention.

XX Sequence 12 AA;

Query Match 37.5%; Score 33; DB 21; Length 12;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSGNCG 9
Db 4 PSGDCG 9

RESULT 9
AAU16769
ID AAU16769 standard; Peptide; 12 AA.
XX
AC AAU16769;

XX 07-NOV-2001 (first entry)
XX
DE Peptide EED18/47/48 derived from Cepsilon2 region of human IgE.
XX
KW Human; linkage technology; conjugated compound; carrier vehicle;
KW epitope; Cepsilon2; Cepsilon3; Cepsilon4; immunoglobulin E;
KW IgE mediated disease; antibody response.

XX Homo sapiens.
OS Synthetic.

XX WO200145745-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-GB04935.

XX 21-DEC-1999; 99GB-0030233.

PR 22-FEB-2000; 2000GB-0004096.

PR 22-AUG-2000; 2000GB-0020707.

PR 22-AUG-2000; 2000GB-0020708.

XX (ACAM-) ACAMBIS RES LTD.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Flinn N, Johnson T;

XX WPI; 2001-521967/57.

XX A linkage comprising an immunogenic conjugate useful treatment of IgE
PT mediated diseases -

XX Example 4; Page 23; 48pp; English.

XX The present invention relates to linkage methodology for use in the
CC conjugation of compounds (e.g. peptides) to carrier vehicles
CC (e.g. macromolecules, polymers, dendrimers, proteins) to produce a
CC biological and immunological constructs. The invention provides a
CC method for linking an epitope (e.g. a peptide) to a carrier (e.g. a
CC protein) for use in a pharmaceutical composition or a vaccine. The
CC invention describes peptides derived from or mimotopes of the
CC Cepsilon2, Cepsilon3 or Cepsilon4 regions of human immunoglobulin E
CC (IgE) which are used to produce conjugated compounds. The compounds or
CC compositions of the invention are useful in the manufacture of a
CC medicament for the treatment of IgE mediated diseases. The invention
CC allows for controlled conjugation of a peptide epitope (antigen) to a
CC protein so as to form an immunogenic conjugate which may be able to
CC raise a protective antibody response in an animal or human patient.
CC AAU16632-AAU16913 represent peptide sequences derived from or mimotopes of
CC the Cepsilon2/Cepsilon3/Cepsilon4 region of human IgE.

XX Sequence 12 AA;

Query Match 37.5%; Score 33; DB 22; Length 12;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSGNCG 9
Db 4 PSGDCG 9

RESULT 10
ABJ00390
ID ABJ00390 standard; Peptide; 12 AA.

XX AC ABJ00390;

XX 02-SEP-2002 (first entry)

XX Human IgE cyclic immunogenic peptide SEQ ID NO: 174.

KW Immunogen; human; IgE; immunoglobulin E; allergy; thio-ether linkage;
KW vaccine; antiallergic; cyclic.
XX
CS Homo sapiens.
XX
PN WC200216409-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-EP09576.
XX
XX 22-AUG-2000; 2000GB-0020717.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA (PEPT-) PEPTIDE THERAPEUTICS LTD.
XX
PI Friede M, Mason S, Turneli WG, Vinals Bassols YC;
XX WPI; 2002-489648/52.
DR
XX
XX Conjugate for use in vaccine for treatment of allergy, comprises
PT disulfide bridge cyclized peptide and immunogenic carrier -
PT
PS Claim 4; Page 13; 45pp; English.
XX
CC The present invention relates to conjugates suitable for use in vaccines,
CC where the conjugate comprises a disulphide bridge cyclised peptide and an
CC immunogenic carrier. The vaccines can be used in the treatment of
CC allergies. The present sequence is a cyclic peptide immunogen derived
CC from human immunoglobulin E (IgE) suitable for use in the invention.
XX
SQ Sequence 12 AA;

Query Match 37.5%; Score 33; DB 23; Length 12;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSGNCG 9
|||:|
DB 4 PSGDCG 9

RESULT 11
AAW23509
ID AAW23509 standard; peptide; 15 AA.
XX
AC AAW23509;
XX
DT 18-SEP-1997 (first entry)
XX
DE Purified cis-9,10-octadecenoamidase internal amino acid fragment.
XX
KW Coase; cis-9,10-octadecenoamide; oleic acid; catalysis;
KW affinity chromatography; electric chromatography;
KW gel filtration chromatography; ion exchange chromatography;
KW partition chromatography; fatty acid primary amide; sleep-inducing;
KW inhibitor; soporific; rat.
XX
OS Rattus rattus.
XX
PN WC9641869-A1.
XX
XX 27-DEC-1996.
XX
PF 12-JUN-1996; 96WO-US10435.
XX
PR 12-JUN-1995; 95US-0489535.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Cravatt BF, Gilula NB, Lerner RA;
XX WPI; 1997-065456/06.
DR

XX Purified cis-9,10-octadecenoamidase - useful for hydrolysing
PT sleep-inducing fatty acid primary amide(s), and identifying
PT inhibitors
XX
PS Claim 1; Page 78; 101pp; English.
XX
CC A purified form of cis-9,10-octadecenoamidase (COase) has been
CC obtained by a chromatographic methodology selected from affinity,
CC electric, gel filtration, ion exchange and partition chromatography.
CC The COase is characterised by enzymic activity for catalysing the
CC conversion of cis-9,10-octadecenoamide (CO) to oleic acid and by the
CC inclusion of an amino acid sequence fragment from rat liver COase.
CC The present sequence represents a specifically claimed example of such
CC an amino acid fragment. This fragment is found at residues 31 to 45 of
CC rat liver COase as shown in AAW10465. The COase can be used to catalyse
CC the hydrolysis of fatty acid primary amides, which have sleep-inducing
CC activity. The COase can also be used to identify inhibitors of the
CC COase activity.
XX
SQ Sequence 15 AA;

Query Match 37.5%; Score 33; DB 18; Length 15;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RYPSGNCGL 10
|:|:|
DB 1 RFPSPAFQGI 9

RESULT 12
AAW57788
ID AAW57788 standard; Peptide; 15 AA.
XX
AC AAW57788;
XX
DT 12-OCT-1998 (first entry)
XX
DE Fatty acid amide hydrolase peptide (e).
XX
KW Cis-9,10-octadecenoamidase; fatty acid amide hydrolase; FAAH;
KW cleamide hydrolase; soporific; sleep.
XX
OS Mammalia.
XX
PN WC9820119-A1.
XX
XX 14-MAY-1998.
XX
PF 04-NOV-1997; 97WO-US20395.
XX
PR 04-NOV-1996; 96US-0743168.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Cravatt BF, Gilula NB, Lerner RA;
XX WPI; 1998-286935/25.
DR
XX
XX New fatty acid amide hydrolase hydrolysing soporific unsaturated
PT amide(s) - useful for, e.g. studying, and potentially developing
PT agents for modulating sleep processes
XX
PS Claim 5; Page 111; 151pp; English.
XX
CC Fatty acid amide hydrolases (FAAHs) of the invention are
CC characterised by inclusion of an amino acid sequence selected from
CC a group of 28 sequences (see AAW57784-811). These FAAHs can hydrolyse
CC cis-9,10-octadecenoamide, anandamide (arachidoyl ethanolamine), and
CC myristic, palmitic or stearic amides. Also new are: (1) inhibiting
CC hydrolysis of fatty acid primary amides catalysed by FAAH by
CC treatment with an FAAH inhibitor; (2) a method of screening for an

CC FAAH inhibitor; (3) the FAAH inhibitor of formula
CC CF3CC(CH2)7-CH=CH(cis)-(CH2)7Me1; and (4) nucleic acid encoding FAAH
CC or parts of it. FAAH catalyses conversion of fatty acid primary
CC amides, particularly those in which the alkyl group has a cis
CC unsaturation. These amides induce sleep, so FAAH can be used to
CC study processes in which they are involved and also to develop
CC agents for modulating sleep.
XX
SQ Sequence 15 AA;

Query Match 37.5%; Score 33; DB 19; Length 15;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RYPSGNCGI 10
|:|:|
Db 1 RFPSAFCGI 9

RESULT 13
AAG96214
ID AAG96214 standard; Peptide; 10 AA.
XX
AC AAG96214;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 2408.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
FN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB04776.
XX
PR 13-DEC-1999; 99GB-C029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides
PT to proteins encoded by genes of the human genome, useful in an assay
PT for screening and identifying of one or more novel peptides which are
PT drug candidates or pro-drugs -
XX
FS Example 4; Page 386; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands
CC generated from the human genome. The complementary peptides
CC interact with their relevant target proteins encoded in the human
CC genome. They can be used as reagents in drug discovery and as lead
CC ligands to facilitate drug design and development. The present
CC sequence is a complementary peptide provided in the specification.
XX
SQ Sequence 10 AA;

Query Match 35.2%; Score 31; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SGNCG 9
|:|:|
Db 6 SGNCG 10

RESULT 14

AAP40328
ID AAP40328 standard; peptide; 15 AA.
XX
AC AAP40328;
XX
DT 16-AUG-2002 (updated)
DT 30-JAN-1992 (first entry)
XX
DE Sequence of conotoxin peptide GIA.
XX
KW Acetylcholine receptor; reversible immobilisation;
KW synaptic transmission inhibitor.
XX
OS Conus geographus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 2..7
FT Modified-site 3
FT /label= Cys-S(acetamido-methyl)
FT Modified-site 13
FT /label= Cys-S(acetamido-methyl)
FT Modified-site 15
FT /label= Lys-NH2
XX
UNUS4447356-A.
XX
PD 08-MAY-1984.
XX
PF 04-JUN-1982; 82US-0385125.
XX
PR 04-JUN-1982; 82US-0385125.
PR 17-APR-1981; 81US-0255237.
XX
PA (OLIV/) OLIVERA B M.
XX
PI Olivera BM, Cruz LC, Gray WR, Rivier JEP;
XX
DR WPI; 1984-133757/21.
XX
PT Conotoxin peptide(s) - useful for reversible immobilisation of
PT muscles and for detecting acetylcholine receptors
XX
PS Claim 5; column 2; 10pp; English.
XX
CC The peptides of the invention are potent inhibitors of synaptic
CC transmission at the neuromuscular junction while lacking inhibition
CC of either nerve or muscle action potential propagation. Their action
CC is freely reversible on dilution or removal of the peptides from the
CC affected muscle. The peptides are useful for reversible
CC immobilisation of a muscle or gp. of muscles in man and other
CC vertebrates and they can be used for detection and measurement of
CC acetylcholine receptors.
CC (Updated on 16-AUG-2002 to add missing OS field.)
XX
SQ Sequence 15 AA;

Query Match 35.2%; Score 31; DB 5; Length 15;
Best Local Similarity 62.5%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYYSSG 15
|:|:|:|
Db 7 CGRHYSCG 14

RESULT 15
AAR75265
ID AAR75265 standard; peptide; 15 AA.
XX
AC AAR75265;
XX
DT 21-DEC-1995 (first entry)

XX Alpha-conotoxin G1A peptide.
DE Alpha conotoxin; inhibit; neuromuscular; synapse; signal transmission.
XX
KW Alpha conotoxin; inhibit; neuromuscular; synapse; signal transmission.
XX
OS Conus geographus.
XX
PN WO9511256-A1.
XX
XX 27-APR-1995.
XX
PF 19-OCT-1994; 94WO-US11927.
XX
PR 19-OCT-1993; 93US-0137900.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI Cruz LJ, Hillyard DR, McIntosh JM, Olivera BM, Santos AD;
XX
DR WPI; 1995-170189/22.
XX
XX New A-lineage conotoxin peptide(s) - which inhibit synaptic
PT transmission at the neuromuscular junction or are active against
PT potassium or sodium channels
XX
PS Disclosure; Page 4; 66pp; English.
XX
CC The kappa-conotoxin, alpha conotoxin and alpha-like conotoxin
CC peptides all belong to a group of peptides known as the A-lineage
CC conotoxin peptides. The A lineage conotoxin peptides have a wide
CC variety of pharmacological uses. The A-lineage conotoxin peptides
CC claimed (AAR75264-R75293) are useful for the inhibition of synaptic
CC transmission at neuromuscular junctions by blocking nicotinic acetyl
CC choline receptors and they also have activity against voltage-gated Na
CC and K channels.
XX
SQ Sequence 15 AA;
Query Match 35.2%; Score 31; DB 16; Length 15;
Best Local Similarity 62.5%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 8 CGLYSSG 15
Db 7 CGRHYSYG 14
Search completed: November 5, 2003, 16:57:26
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 16:58:41 ; Search time 28 Seconds
(without alignments)
92.009 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 112152

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOXB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOXB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOXB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOXB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOXB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOXB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOXB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOXB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOXB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOXB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOXB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description:
1	88	100.0	15	11	US-09-902-563-18 Sequence 18, Appl
2	88	100.0	15	15	US-10-096-255-18 Sequence 18, Appl
3	37	42.0	14	15	US-10-056-884-11 Sequence 11, Appl
4	35	39.8	10	9	US-09-819-308-25 Sequence 25, Appl
5	33	37.5	9	12	US-09-793-451-52 Sequence 52, Appl
6	33	37.5	9	12	US-09-793-451-518 Sequence 518, App
7	33	37.5	9	12	US-09-793-451-622 Sequence 622, App
8	33	37.5	9	12	US-09-793-451-640 Sequence 640, App
9	33	37.5	9	12	US-10-283-722-52 Sequence 52, Appl
10	33	37.5	9	12	US-10-283-722-518 Sequence 518, App
11	33	37.5	9	12	US-10-283-722-622 Sequence 622, App
12	33	37.5	9	12	US-10-283-722-640 Sequence 640, App
13	33	37.5	10	12	US-09-793-451-683 Sequence 683, App
14	33	37.5	10	12	US-10-283-722-683 Sequence 683, App
15	31	35.2	10	11	US-09-572-404B-2408 Sequence 2408, Ap

16	30	34.1	10	11	US-09-572-404B-490 Sequence 490, App
17	30	34.1	12	12	US-10-218-906-20 Sequence 20, Appl
18	30	34.1	14	12	US-10-125-187-16 Sequence 16, Appl
19	30	34.1	14	12	US-10-125-187-17 Sequence 17, Appl
20	30	34.1	14	12	US-10-125-187-50 Sequence 50, Appl
21	30	34.1	14	12	US-10-125-187-51 Sequence 51, Appl
22	29	33.0	10	12	US-09-793-451-572 Sequence 572, App
23	29	33.0	10	12	US-09-793-451-677 Sequence 677, App
24	29	33.0	10	12	US-10-283-722-572 Sequence 572, App
25	29	33.0	10	12	US-10-283-722-677 Sequence 677, App
26	28.5	32.4	10	7	US-08-344-824-234 Sequence 234, App
27	28	31.8	6	15	US-10-131-543-3 Sequence 3, Appli
28	28	31.8	6	15	US-10-131-546-3 Sequence 3, Appli
29	28	31.8	6	15	US-10-131-346-3 Sequence 3, Appli
30	28	31.8	11	11	US-09-924-993-6 Sequence 6, Appli
31	28	31.8	15	15	US-10-161-499-55 Sequence 55, Appl
32	27	30.7	7	11	US-09-792-286-221 Sequence 221, App
33	27	30.7	7	11	US-09-792-286-225 Sequence 225, App
34	27	30.7	9	9	US-09-832-723-80 Sequence 80, Appl
35	27	30.7	3	12	US-10-303-331-80 Sequence 80, Appl
36	27	30.7	10	11	US-09-572-404B-488 Sequence 488, App
37	27	30.7	10	11	US-09-572-404B-3542 Sequence 3542, Ap
38	27	30.7	10	11	US-09-572-404B-3544 Sequence 3544, Ap
39	27	30.7	10	12	US-09-572-270A-581 Sequence 581, App
40	27	30.7	11	11	US-09-852-910-48 Sequence 48, Appl
41	27	30.7	14	11	US-09-924-993-7 Sequence 7, Appli
42	27	30.7	14	12	US-10-125-187-18 Sequence 18, Appl
43	27	30.7	14	12	US-10-125-187-52 Sequence 52, Appl
44	27	30.7	14	15	US-10-219-834-97 Sequence 97, Appl
45	27	30.7	15	14	US-10-086-623-25 Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-902-563-18
; Sequence 18, Application US/09902563
; Publication No. US20030099654A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-37
; CURRENT APPLICATION NUMBER: US/09/902,563
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: US 09/442,143
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-902-563-18

Query Match 100.0%; Score 88; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRYPSGNCGLYSSG 15
|||
Db 1 DRYPSGNCGLYSSG 15

RESULT 2
US-10-096-255-18
; Sequence 18, Application US/10096255
; Publication No. US20030103974A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; APPLICANT: Clark, David A.
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-52

; CURRENT APPLICATION NUMBER: US/10/096,255
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-096-255-18

Query Match 100.0%; Score 82; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
DB 1 DRYPSGNCGLYSSG 15

RESULT 3

US-10-056-884-11
; Sequence 11; Application US/10056884
; Publication No. US20030032786A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN POTASSIUM CHANNEL BETA-SUB
; FILE OF INVENTION: K-betaM2
; FILE REFERENCE: D0076 N2
; CURRENT APPLICATION NUMBER: US/10/056,884
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,872
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/269,794
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-056-884-11

Query Match 42.0%; Score 37; DB 15; Length 14;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12
DB 4 SGNCSRY 11

RESULT 4

US-09-819-308-25
; Sequence 25; Application JS/09819308
; Patent No. US20020019040A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020019040Aleborn, Mathieu
; APPLICANT: Danen-van Corschot, Astrid
; APPLICANT: Rohr, Jennifer
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN
; FILE REFERENCE: 2906-4820US
; CURRENT APPLICATION NUMBER: US/09/919,308
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Sequence homology analysis of AAP-5

US-09-819-308-25

Query Match 39.8%; Score 35; DB 9; Length 10;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGJYSS 14
DB 3 CGKYSS 9

RESULT 5

US-09-793-451-52
; Sequence 52; Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-793-451-52

Query Match 37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGJY 12
DB 2 SGRGGLY 9

RESULT 6

US-09-793-451-518
; Sequence 518; Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 518
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-518
```

```
Query Match          37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches      6; Conservative      0; Mismatches      2; Indels      0; Gaps      0;
```

```
Cy      5 SGNCGLYY 12
      |||||
Db      1 SGRCGGLGY 8
```

RESULT 7

```
US-09-793-451-622
; Sequence 622, Application US/C9793451
; Publication No. US20030157597A1
```

GENERAL INFORMATION:

```
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-622
```

```
Query Match          37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches      6; Conservative      0; Mismatches      2; Indels      0; Gaps      0;
```

```
Cy      5 SGNCGLYY 12
      |||||
Db      2 SGRCGGLGY 9
```

RESULT 8

```
US-09-793-451-640
; Sequence 640, Application US/C9793451
; Publication No. US20030157597A1
```

GENERAL INFORMATION:

```
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
```

```
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 640
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-640
```

```
Query Match          37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches      6; Conservative      0; Mismatches      2; Indels      0; Gaps      0;
```

```
Cy      5 SGNCGLYY 12
      |||||
Db      1 SGRCGGLGY 8
```

RESULT 9

```
US-10-283-722-52
; Sequence 52, Application US/10283722
; Publication No. US20030194407A1
```

GENERAL INFORMATION:

```
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/793,451
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-52
```

```
Query Match          37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches      6; Conservative      0; Mismatches      2; Indels      0; Gaps      0;
```

```
Cy      5 SGNCGLYY 12
      |||||
Db      2 SGRCGGLGY 9
```

RESULT 10

```
US-10-283-722-518
; Sequence 518, Application US/10283722
; Publication No. US20030194407A1
```

GENERAL INFORMATION:

```
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
```

```
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; PRIOR FILING DATE: 2003-02-03
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 518
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-518
```

```
Query Match: 37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 2; Gaps 0;
```

```
QY 5 SGNCGLY 12
Db 1 SGRCGLY 8
```

```
RESULT 11
US-10-283-722-622
; Sequence 622, Application US/10283722
; Publication No. US200301944C7A;
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; PRIOR FILING DATE: 2003-02-03
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-622
```

```
Query Match: 37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 2; Gaps 0;
```

```
QY 5 SGNCGLY 12
Db 2 SGRCGLY 9
```

```
RESULT 12
US-10-283-722-640
```

```
; Sequence 640, Application US/10283722
; Publication No. US200301944C7A;
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; PRIOR FILING DATE: 2003-02-03
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 640
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-640
```

```
Query Match: 37.5%; Score 33; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 5.8e+05;
Matches 6; Conservative 0; Mismatches 2; Indels 2; Gaps 0;
```

```
QY 5 SGNCGLY 12
Db 1 SGRCGLY 8
```

```
RESULT 13
US-09-793-451-683
; Sequence 693, Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2001-02-26
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 683
; LENGTH: 10
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-683
```

```
Query Match: 37.5%; Score 33; DB 12; Length 10;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 2; Indels 2; Gaps 0;
```

Qy 5 SGNCGLYY 12
|||
Db 2 SGRCGJGY 9

Db 6 SGNCG 10

Search completed: November 5, 2003, 17:03:48
Job time : 28 secs

RESULT 14

US-10-283-722-683
; Sequence 683, Application US/10283722
; Publication No. US20030194407A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Chailita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/10/283,722
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/793,451
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 683
; LENGTH: 10
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-283-722-683

Query Match 37.5%; Score 33; DB 12; Length 10;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SGNCGLYY 12
|||
Db 2 SGRCGJGY 9

RESULT 15

US-09-572-404B-2408
; Sequence 2408, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 2408
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in GATA1 OR GFI OR ERYF1 at 338-347 and may inte
; OTHER INFORMATION: with Sequence 2407 in this patent.
US-09-572-404B-2408

Query Match 35.2%; Score 31; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 SGNCG 9
:|:|:|

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 5, 2003, 16:56:41 ; Search time 22 Seconds
(without alignments)
28.848 Million cell updates/sec

Title: US-09-902-563-18
Perfect score: 88
Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310358 residues

Total number of hits satisfying chosen parameters: 128097

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	15	4 US-09-442-143A-18	Sequence 18, Appl
2	33	37.5	15	3 US-08-743-168B-9	Sequence 9, Appli
3	33	37.5	15	5 PCT-US96-10435-9	Sequence 9, Appli
4	31	35.2	15	1 US-08-137-800-5	Sequence 5, Appli
5	31	35.2	15	1 US-08-477-383-5	Sequence 5, Appli
6	31	35.2	15	1 US-08-487-174-5	Sequence 5, Appli
7	31	35.2	15	1 US-08-480-750-5	Sequence 5, Appli
8	31	35.2	15	5 PCT-US96-07962-2	Sequence 2, Appli
9	30	34.1	13	1 US-07-923-724-57	Sequence 57, Appl
10	30	34.1	13	2 US-08-609-426A-57	Sequence 57, Appl
11	30	34.1	13	2 US-08-374-652C-44	Sequence 44, Appl
12	30	34.1	15	2 US-08-553-257A-42	Sequence 42, Appl
13	30	34.1	15	4 US-09-441-992-42	Sequence 42, Appl
14	29	33.0	15	2 US-08-432-871C-91	Sequence 91, Appl
15	29	33.0	15	4 US-09-270-956-91	Sequence 91, Appl
16	28.5	32.4	10	2 US-08-318-856A-72	Sequence 72, Appl
17	28	31.8	7	4 US-08-492-411A-20	Sequence 20, Appl
18	28	31.8	11	1 US-08-466-468-6	Sequence 6, Appli
19	28	31.8	11	3 US-08-468-408-6	Sequence 6, Appli
20	28	31.8	11	3 US-08-937-228-6	Sequence 6, Appli
21	28	31.8	11	4 US-08-492-411A-33	Sequence 33, Appl
22	28	31.8	11	4 US-08-421-583-6	Sequence 6, Appli
23	28	31.8	11	4 US-09-639-242A-6	Sequence 6, Appli
24	28	31.8	12	1 US-08-190-786A-21	Sequence 21, Appl
25	28	31.8	12	1 US-08-383-474B-26	Sequence 26, Appl
26	28	31.8	12	1 US-08-465-391A-21	Sequence 21, Appl
27	28	31.8	12	2 US-08-464-538B-21	Sequence 21, Appl

28	28	31.8	12	2 US-08-463-076E-65	Sequence 65, Appl
29	28	31.8	12	3 US-08-737-841-13	Sequence 13, Appl
30	28	31.8	13	1 US-08-089-994A-20	Sequence 20, Appl
31	28	31.8	13	5 PCT-US94-07605-20	Sequence 20, Appl
32	28	31.8	15	2 US-08-432-871C-100	Sequence 100, App
33	28	31.8	15	4 US-09-347-504-55	Sequence 55, App
34	28	31.8	15	4 US-09-270-956-100	Sequence 100, App
35	27	30.7	13	3 US-08-881-037-88	Sequence 88, Appl
36	27	30.7	14	1 US-08-466-468-7	Sequence 7, Appli
37	27	30.7	14	3 US-08-468-408-7	Sequence 7, Appli
38	27	30.7	14	3 US-08-937-228-7	Sequence 7, Appli
39	27	30.7	14	4 US-08-421-583-7	Sequence 7, Appli
40	27	30.7	14	4 US-09-639-242A-7	Sequence 7, Appli
41	27	30.7	15	2 US-08-432-871C-80	Sequence 80, Appl
42	27	30.7	15	2 US-08-432-871C-83	Sequence 83, Appl
43	27	30.7	15	3 US-08-469-141A-14	Sequence 14, Appl
44	27	30.7	15	4 US-09-347-504-68	Sequence 68, Appl
45	27	30.7	15	4 US-09-347-504-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1

US-09-442-143A-18

; Sequence 18, Application US/09442143A

; Patent No. 6403C89

; GENERAL INFORMATION:

; APPLICANT: Levy, Gary

; APPLICANT: Clark, David A.

; TITLE OF INVENTION: Methods of Modulating Immune Coagulation

; FILE REFERENCE: 9579-14

; CURRENT APPLICATION NUMBER: US/09/442,143A

; CURRENT FILING DATE: 1999-11-15

; PRIOR APPLICATION NUMBER: US 60/046,537

; PRIOR FILING DATE: 1997-05-17

; PRIOR APPLICATION NUMBER: US 60/061,684

; PRIOR FILING DATE: 1997-10-10

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 18

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-442-143A-18

Query Match 100.0%; Score 88; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. NO. 4.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy : DRYPSGNCGLYSSG 15
||| ||| ||| ||| |||
Db 1 DRYPSGNCGLYSSG 15

RESULT 2

US-08-743-168B-9

; Sequence 9, Application US/08743168B

; Patent No. 6271015

; GENERAL INFORMATION:

; APPLICANT: Gilula, No. 6271015ton B

; APPLICANT: Cravatt, Benjamin F

; APPLICANT: Lerner, Richard A

; TITLE OF INVENTION: FATTY-ACID AMIDE HYDROLASE

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Scripps Research Institute

; STREET: 10550 No. 6271015th Torrey Pines Road

; CITY: La Jolla

; STATE: California

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/743,168B
/ FILING DATE: 04-NOV-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/489,535
/ FILING DATE: 12-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fitting, Thomas
/ REGISTRATION NUMBER: 34,163
/ REFERENCE/DOCKET NUMBER: TSR: 485.2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 784-2937
/ TELEFAX: (619) 784-9399
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: internal
/ US-08-743-168B-9

Query Match 37.5%; Score 33; DB 3; Length 15;
Best Local Similarity 55.6%; Pred. No. 70;
Matches 5; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 2 RYPSGNCGL 10
|:|:| |:
Db 1 RFPSPFCGI 9

RESULT 3
PCT-US96-10435-9
/ Sequence 9, Application: PC/TJS9610435
/ GENERAL INFORMATION:
/ APPLICANT: The Scripps Research Institute
/ TITLE OF INVENTION: CIS-9, 10-OCTADECENOAMIDASE
/ NUMBER OF SEQUENCES: 32
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/10435
/ FILING DATE: 12-JUN-1996
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/489,535
/ FILING DATE: 12-JUN-1995
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ FRAGMENT TYPE: internal
/ PCT-US96-10435-9

Query Match 37.5%; Score 33; DB 5; Length 15;
Best Local Similarity 55.6%; Pred. No. 70;
Matches 5; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 2 RYPSGNCGL 10
|:|:| |:
Db 1 RFPSPFCGI 9

RESULT 4
US-08-137-800-5
/ Sequence 5, Application US/08137800
/ Patent No. 5514774
/ GENERAL INFORMATION:
/ APPLICANT: Olivera, Baldomero M.
/ APPLICANT: Cruz, Lourdes J.
/ APPLICANT: Hillyard, David R.
/ APPLICANT: McIntosh, J. Michael
/ APPLICANT: Santos, Ameurfin D.
/ TITLE OF INVENTION: Conotoxin Peptides
/ NUMBER OF SEQUENCES: 53
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Venable, Baetjer, Howard & Civiletti
/ STREET: 1201 New York Avenue N.W., Suite 1000
/ City: Washington
/ STATE: DC
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WordPerfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/137,800
/ FILING DATE: 19-OCT-1993
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ihnen, Jeffrey L.
/ REGISTRATION NUMBER: 28,957
/ REFERENCE/DOCKET NUMBER: 24260-1C4763
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-962-4810
/ TELEFAX: 202-962-8300
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Conus geographus
/ US-08-137-800-5

Query Match 35.2%; Score 31; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 14e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYSSG 15
|:|:| |:
Db 7 CRRHSCG 14

RESULT 5
US-08-477-383-5
/ Sequence 5, Application US/08477383
/ Patent No. 5589340
/ GENERAL INFORMATION:
/ APPLICANT: Olivera, Baldomero M.
/ APPLICANT: Cruz, Lourdes J.
/ APPLICANT: Hillyard, David R.
/ APPLICANT: Macintosh, J. Michael
/ APPLICANT: Santos, Ameurfin S.
/ TITLE OF INVENTION: Conotoxin Peptides
/ NUMBER OF SEQUENCES: 59
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Venable, Baetjer, Howard & Civiletti
/ STREET: 1201 New York Avenue, N.W., Suite 1000
/ City: Washington
/ STATE: DC
/ COUNTRY: U.S.A.

ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,383
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus geographus
US-08-477-383-5

Query Match 35.2%; Score 31; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 8 CGLYSSG 15
||:||||
Db 7 CGRHYSCG 14

RESULT 6
US-08-487-174-5
Sequence 5, Application US/08487174
Patent No. 5595972
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,174
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus geographus
US-08-487-174-5

Query Match 35.2%; Score 31; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 8 CGLYSSG 15
||:||||
Db 7 CGRHYSCG 14

RESULT 7
US-08-480-750-5
Sequence 5, Application US/08480750
Patent No. 5633347
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hillyard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,750
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810

```

; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus geographus
US-08-480-750-5

Query Match 35.2%; Score 31; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CGLYYSSG 15
Db 7 CGRHYSCG 14

RESULT 8
PCT-US96-07962-2
; Sequence 2, Application: PC/TJS9607962
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Use of Conotoxin Peptides U002 and M11
; TITLE OF INVENTION: for Treating or Detecting Small-Cell Lung Carcinoma
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-WINDOWS
; SOFTWARE: Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07962
; FILING DATE: 04-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,174
; FILING DATE: 07-JUN-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus geographus
PCT-US96-07962-2

Query Match 35.2%; Score 31; DB 5; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CGLYYSSG 15
Db 7 CGRHYSCG 14
```

```

RESULT 9
US-07-923-724-57
; Sequence 57, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerström, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: both
US-07-923-724-57

Query Match 34.1%; Score 30; DB 1; Length 13;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DRYPSGNCG 9
Db 4 ERYPSPSAG 12

RESULT 10
US-08-609-426A-57
; Sequence 57, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
```

```
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstrom, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DCS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/609,426A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant B.
; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: both
;
US-08-609-426A-57

Query Match 34.1%; Score 30; DB 2; Length 13;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSGNCG 9
DB 4 ERYSPSAG 12

RESULT 11
US-08-374-652C-44
; Sequence 44, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
```

```
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
; TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DCS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,652C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07058
; FILING DATE: 27-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925,401
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT B.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1050.071001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
;
US-08-374-652C-44

Query Match 34.1%; Score 30; DB 2; Length 13;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSGNCG 9
DB 4 ERYSPSAG 12

RESULT 12
US-08-553-257A-42
; Sequence 42, Application US/08553257A
; Patent No. 5994083
; GENERAL INFORMATION:
; APPLICANT: ISTITUTO DI RICERCHE DI BIOLOGIA
; APPLICANT: MOLECOLARE P. ANGELETTI S.p.A.
; APPLICANT: FELICI, Franco
; APPLICANT: LUZZAGO, Alessandra
; APPLICANT: NICOSIA, Alfredo
; APPLICANT: MONACI, Paolo
; APPLICANT: CORTESE, Riccardo
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF IMMUNOGENS
; TITLE OF INVENTION: OR DIAGNOSTIC REAGENTS, AND IMMUNOGENS OR
; TITLE OF INVENTION: DIAGNOSTIC REAGENTS THEREBY OBTAINABLE
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
```

STREET: 419 Seventh Street N.W. Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,257A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IT94/00054
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: RM93A000301
FILING DATE: 11-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: FELICI=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: recombinant protein
HYPOTHETICAL: yes
FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
LIBRARY: of recombinant peptides on phage
CLONE: phagic
FEATURE:
NAME/KEY: polypeptide
IDENTIFICATION METHOD: selection with specific antibodies
US-08-553-257A-42
Query Match 34.1%; Score 30; DB 2; Length 15;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 PSGNCG 9
Db 9 PSGACG 14
RESULT 13
US-09-441-992-42
Sequence 42, Application US/09441992
Patent No. 6541210
GENERAL INFORMATION:
APPLICANT: ISTITUTO DI RICERCA DI BIOLOGIA
MOLECOLARE P. ANGELETTI S.P.A.
FELICI, Franco
LUZZAGO, Alessandra
NICOSIA, Alfredo
MONACI, Paolo
CORTESE, Riccardo
TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF IMMUNOGENS
OR DIAGNOSTIC REAGENTS, AND IMMUNOGENS OR
DIAGNOSTIC REAGENTS THEREBY OBTAINABLE
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W. Ste. 300
CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/441,992
FILING DATE: 18-No. 6541210-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/553,257
FILING DATE: <Unknown>
APPLICATION NUMBER: RM93A000301
FILING DATE: 11-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: FELICI=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: recombinant protein
HYPOTHETICAL: yes
FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
LIBRARY: of recombinant peptides on phage
CLONE: phagic
FEATURE
(ix)
(A) NAME: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-09-441-992-42
Query Match 34.1%; Score 30; DB 4; Length 15;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 PSGNCG 9
Db 9 PSGACG 14
RESULT 14
US-08-432-871C-91
Sequence 91, Application US/08432871C
Patent No. 5877010
GENERAL INFORMATION:
APPLICANT: Loeb, Lawrence A.
APPLICANT: Black, Margaret S.
TITLE OF INVENTION: THYMIDINE KINASE MUTANTS
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,871C

Search completed: November 5, 2003, 16:59:42
Job time : 22 secs

FILING DATE: 02-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.409C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-432-871C-91

Query Match 33.0%; Score 29; DB 2; Length 15;
Best Local Similarity 45.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 DRYPSGNCGLY 11
DB 4 DRHPIGQTSCY 14

RESULT 15
US-09-270-956-91
Sequence 91, Application US/09270956
Patent No. 6451571
GENERAL INFORMATION:
APPLICANT: Loeb, Lawrence A.
APPLICANT: Black, Margaret E.
TITLE OF INVENTION: THYMIDINE KINASE MUTANTS
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270,956
FILING DATE: 17-MAR-1999
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.409C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-270-956-91

Query Match 33.0%; Score 29; DB 4; Length 15;
Best Local Similarity 45.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 DRYPSGNCGLY 11
DB 4 DRHPIGQTSCY 14